```
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-143-822

Alignment Scores:
Pred. No.:              4.07e-72        Length:         675
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-759-143-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 2
US-09-780-669-822
; Sequence 822, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C24
;  CURRENT APPLICATION NUMBER: US/09/780,669
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS:  943
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-780-669-822

Alignment Scores:
Pred. No.:              4.07e-72        Length:         675
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-780-669-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 3
US-09-822-827-822
; Sequence 822, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C1
;  CURRENT APPLICATION NUMBER: US/09/822,827
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-822-827-822

Alignment Scores:
```

```
Pred. No.:              4.07e-72        Length:         675
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-822-827-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 4
US-09-895-793-822
; Sequence 822, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-793-822

Alignment Scores:
Pred. No.:              4.07e-72        Length:         675
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-895-793-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 5
US-09-895-814-822
; Sequence 822, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C26
;  CURRENT APPLICATION NUMBER: US/09/895,814
```

; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-822

Alignment Scores:

Pred. No.:	4.07e-72	Length:	675
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-895-814-822 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 6
US-10-012-896-822
; Sequence 822, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-822

Alignment Scores:

Pred. No.:	4.07e-72	Length:	675
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-09-684-215B-18 (1-128) x US-10-012-896-822 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 7
US-10-144-678A-822
; Sequence 822, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.

```
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A. W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals y de Bassols, Carlota
;  APPLICANT:  Foy, Teresa M.
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Deng, Ta
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C28
;  CURRENT APPLICATION NUMBER: US/10/144,678A
;  CURRENT FILING DATE:  2002-08-12
;  NUMBER OF SEQ ID NOS: 1033
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-144-678A-822

Alignment Scores:
Pred. No.:                4.07e-72        Length:         675
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       15              Gaps:           0

US-09-684-215B-18 (1-128) x US-10-144-678A-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 8
US-10-294-025-822
; Sequence 822, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Kalos, Michael D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C29
;  CURRENT APPLICATION NUMBER: US/10/294,025
;  CURRENT FILING DATE:  2002-11-12
;  NUMBER OF SEQ ID NOS: 1038
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-294-025-822

Alignment Scores:
Pred. No.:                4.07e-72        Length:         675
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       15              Gaps:           0

US-09-684-215B-18 (1-128) x US-10-294-025-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 9
US-09-736-457-1862
; Sequence 1862, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1862
;   LENGTH: 822
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

US-09-736-457-1862

Alignment Scores:

Pred. No.:	5.24e-72	Length:	822
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-736-457-1862 (1-822)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 10
US-09-902-941-1862
; Sequence 1862, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-1862

Alignment Scores:

Pred. No.:	5.24e-72	Length:	822
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-902-941-1862 (1-822)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 11
US-09-849-626-1862
; Sequence 1862, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-849-626-1862

Alignment Scores:

Pred. No.:	5.24e-72	Length:	822
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-849-626-1862 (1-822)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81
```

```
Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 12
US-10-283-017-1862
; Sequence 1862, Application US/10283017
; Publication No. US20030211510A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C20
;  CURRENT APPLICATION NUMBER: US/10/283,017
;  CURRENT FILING DATE:  2002-10-28
;  NUMBER OF SEQ ID NOS: 2157
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1862
;   LENGTH: 822
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-283-017-1862

Alignment Scores:
Pred. No.:                5.24e-72        Length:         822
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       13              Gaps:           0

US-09-684-215B-18 (1-128) x US-10-283-017-1862 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 13
US-10-017-754-1862
; Sequence 1862, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C18
;  CURRENT APPLICATION NUMBER: US/10/017,754
;  CURRENT FILING DATE:  2001-10-29
;  NUMBER OF SEQ ID NOS: 2004
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1862
;   LENGTH: 822
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-017-754-1862

Alignment Scores:
Pred. No.:                5.24e-72        Length:         822
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       15              Gaps:           0

US-09-684-215B-18 (1-128) x US-10-017-754-1862 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 14
US-10-113-872-1862
; Sequence 1862, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
;  APPLICANT: Watanabe, Yoshihiro
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C19
;  CURRENT APPLICATION NUMBER: US/10/113,872
;  CURRENT FILING DATE:  2002-03-28
;  NUMBER OF SEQ ID NOS: 2011
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1862
;   LENGTH: 822
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-113-872-1862

Alignment Scores:
Pred. No.:                5.24e-72        Length:         822
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       15              Gaps:           0

US-09-684-215B-18 (1-128) x US-10-113-872-1862 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 15
US-09-878-722-236
; Sequence 236, Application US/09878722
; Publication No. US20020040127A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Clapper, Jonathan
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.524
;  CURRENT APPLICATION NUMBER: US/09/878,722
;  CURRENT FILING DATE:  2001-06-08
;  NUMBER OF SEQ ID NOS: 245
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 236
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-878-722-236

Alignment Scores:
Pred. No.:                5.83e-72        Length:         894
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       13              Gaps:           0

US-09-684-215B-18 (1-128) x US-09-878-722-236 (1-894)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 16
US-09-904-456-236
; Sequence 236, Application US/09904456
; Publication No. US20030017167A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
```

```
;  FILE REFERENCE: 210121.524C1
;  CURRENT APPLICATION NUMBER: US/09/904,456
;  CURRENT FILING DATE:  2001-07-11
;  NUMBER OF SEQ ID NOS: 247
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 236
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-904-456-236

Alignment Scores:
Pred. No.:              5.83e-72        Length:         894
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     13              Gaps:           0

US-09-684-215B-18 (1-128) x US-09-904-456-236 (1-894)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 17
US-09-735-705-353
; Sequence 353, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Fanger, Neil
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C14
;  CURRENT APPLICATION NUMBER: US/09/735,705
;  CURRENT FILING DATE:  2000-12-12
;  NUMBER OF SEQ ID NOS:  419
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 353
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-735-705-353

Alignment Scores:
Pred. No.:              5.88e-72        Length:         900
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-735-705-353 (1-900)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 18
US-09-850-716A-353
; Sequence 353, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
;  APPLICANT: Kalos, Michael D.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Retter, Marc W.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C15
;  CURRENT APPLICATION NUMBER: US/09/850,716A
;  CURRENT FILING DATE:  2001-05-07
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 353
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-850-716A-353

Alignment Scores:
Pred. No.:              5.88e-72        Length:         900
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0
```

US-09-684-215B-18 (1-128) x US-09-850-716A-353 (1-900)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 19
US-09-897-778-353
; Sequence 353, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-353

Alignment Scores:

Pred. No.:	5.88e-72	Length:	900
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-897-778-353 (1-900)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 20
US-10-007-700-353
; Sequence 353, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samual X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-007-700-353

Alignment Scores:

Pred. No.:	5.88e-72	Length:	900
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-09-684-215B-18 (1-128) x US-10-007-700-353 (1-900)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy        41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy        61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy        81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy       101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy       121 ThrLeuAlaGluGlyProProAla 128
             |||||||||||||||||||||||||
Db       382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 21
US-10-117-982-353
; Sequence 353, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
;  APPLICANT: Foy, Teresa M.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Mericle, Barbara
;  APPLICANT:  Spies, Gregory A.
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Tongtong
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C18
;  CURRENT APPLICATION NUMBER: US/10/117,982
;  CURRENT FILING DATE:  2002-04-05
;  NUMBER OF SEQ ID NOS: 484
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 353
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-117-982-353

Alignment Scores:
Pred. No.:                5.88e-72        Length:          900
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       15              Gaps:            0

US-09-684-215B-18 (1-128) x US-10-117-982-353 (1-900)

Qy         1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy        21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy        41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy        61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy        81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy       101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy       121 ThrLeuAlaGluGlyProProAla 128
             |||||||||||||||||||||||||
Db       382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 22
US-10-313-986-353
; Sequence 353, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
;  APPLICANT: Foy, Teresa M.
;  APPLICANT:  McNabb, Andria
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Wang, Tongtong
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C19
;  CURRENT APPLICATION NUMBER: US/10/313,986
;  CURRENT FILING DATE:  2002-12-04
;  NUMBER OF SEQ ID NOS: 560
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 353
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-313-986-353

Alignment Scores:
Pred. No.:                5.88e-72        Length:          900
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       16              Gaps:            0

US-09-684-215B-18 (1-128) x US-10-313-986-353 (1-900)

Qy         1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy        21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy        41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy        61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy        81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy       101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381
```

```
Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 23
US-09-759-143-834
; Sequence 834, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-143-834

Alignment Scores:

Pred. No.:	6.01e-72	Length:	915
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-759-143-834 (1-915)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381
```

```
Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 24
US-09-780-669-834
; Sequence 834, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-780-669-834

Alignment Scores:

Pred. No.:	6.01e-72	Length:	915
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-780-669-834 (1-915)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
                ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 25
US-09-822-827-834
; Sequence 834, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-827-834

Alignment Scores:

Pred. No.:	6.01e-72	Length:	915
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-822-827-834 (1-915)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
                ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 26
US-09-895-793-834
; Sequence 834, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-793-834

Alignment Scores:

Pred. No.:	6.01e-72	Length:	915
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-895-793-834 (1-915)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
                ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 27
US-09-895-814-834
; Sequence 834, Application US/09895814
; Publication No. US20020193296A1

```
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C26
;  CURRENT APPLICATION NUMBER: US/09/895,814
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 990
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
;   LENGTH: 915
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-814-834

Alignment Scores:
Pred. No.:                6.01e-72        Length:          915
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-09-684-215B-18 (1-128) x US-09-895-814-834 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

```
RESULT 28
US-10-012-896-834
; Sequence 834, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Wantanabe, Yoshihiro
;  APPLICANT:  Meagher, Madeleine Joy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C27
;  CURRENT APPLICATION NUMBER: US/10/012,896
;  CURRENT FILING DATE:  2001-12-10
;  NUMBER OF SEQ ID NOS: 1011
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
;   LENGTH: 915
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-012-896-834

Alignment Scores:
Pred. No.:                6.01e-72        Length:          915
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       14              Gaps:            0

US-09-684-215B-18 (1-128) x US-10-012-896-834 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321
```

```
Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 29
US-10-144-678A-834
; Sequence 834, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals y de Bassols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C28
; CURRENT APPLICATION NUMBER: US/10/144,678A
; CURRENT FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 1033
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-144-678A-834

Alignment Scores:

Pred. No.:	6.01e-72	Length:	915
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-18 (1-128) x US-10-144-678A-834 (1-915)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 30
US-10-294-025-834
; Sequence 834, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294,025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-294-025-834

Alignment Scores:

Pred. No.:	6.01e-72	Length:	915
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-18 (1-128) x US-10-294-025-834 (1-915)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

Search completed: April 30, 2004, 05:10:24
Job time : 471.379 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 22:25:27 ; Search time 2507.02 Seconds
(without alignments)
1524.660 Million cell updates/sec

Title: US-09-684-215B-18
Perfect score: 653
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....QTKSGGTRTGNVTLAEGPPA 128

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061305_13200/app_query.fasta_1.1180
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_5167_@runat_29042004_061305_13200 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	145.5	22.3	289	29	U82114	U82114 U82114 orde
c 2	125.5	19.2	1438	28	BH770798	BH770798 LLMGtag54
c 3	123	18.8	959	28	BZ549048	BZ549048 pacs1-60_
c 4	122.5	18.8	726	14	CB679186	CB679186 OSJNEf02F
5	119.5	18.3	590	14	CB925602	CB925602 ABA1_22_F
c 6	117.5	18.0	807	14	CA228900	CA228900 SCQGFL305
c 7	114	17.5	758	13	BQ514888	BQ514888 EST622303
8	106	16.2	603	28	AZ934428	AZ934428 BJ__Ba000
9	104	15.9	765	9	AJ558965	AJ558965 AJ558965
c 10	100.5	15.4	1033	28	BZ561390	BZ561390 pacs2-164
c 11	100	15.3	947	28	BZ549047	BZ549047 pacs1-60_
12	99	15.2	726	28	AQ989479	AQ989479 Rfc00025
c 13	98	15.0	1093	13	BU557763	BU557763 AGENCOURT
14	97	14.9	444	13	CA148171	CA148171 SCEZRZ101
15	96.5	14.8	790	28	BZ577987	BZ577987 msh2_5667
16	96	14.7	546	10	AW285510	AW285510 LG1_241_E
17	96	14.7	551	10	AW285527	AW285527 LG1_241_G
18	95.5	14.6	677	12	BI378928	BI378928 BFLG1_000
19	95	14.5	543	12	BI721127	BI721127 1031054B1
20	95	14.5	673	12	BI956121	BI956121 HVSMEm002
c 21	94.5	14.5	1127	28	BZ560689	BZ560689 pacs2-164
22	94	14.4	982	14	CD080132	CD080132 MA3-9999U
c 23	93.5	14.3	771	29	CG322895	CG322895 OG0DK42TV
c 24	93.5	14.3	787	29	CC634156	CC634156 OGOBL39TV
c 25	93.5	14.3	885	28	BZ567978	BZ567978 pacs2-164
26	92	14.1	703	14	CD383492	CD383492 PTMM08663
c 27	91.5	14.0	600	12	BG143271	BG143271 ia95h02.y
28	91	13.9	701	14	CD204316	CD204316 HS1_7_H09
29	91	13.9	1098	13	BU553475	BU553475 AGENCOURT
c 30	90.5	13.9	631	14	CF512107	CF512107 CAbud0004
c 31	90.5	13.9	728	12	BJ285991	BJ285991 BJ285991
c 32	90	13.8	736	14	CA221763	CA221763 SCEZFL404
c 33	89	13.6	509	10	BE195909	BE195909 HVSMEh009
c 34	88	13.5	655	14	CB921057	CB921057 VVD048E11
c 35	88	13.5	999	12	BM051288	BM051288 603634270
36	87.5	13.4	600	28	CC345457	CC345457 OGQAF70TH
37	87.5	13.4	643	9	AL883551	AL883551 AL883551
38	87.5	13.4	719	28	AZ933900	AZ933900 BJ__Ba000
c 39	87.5	13.4	898	10	BF309796	BF309796 601892449
40	87	13.3	510	13	BQ977384	BQ977384 QHI24M07.
c 41	87	13.3	540	28	BZ893820	BZ893820 HL7_0146
c 42	87	13.3	642	12	BJ283688	BJ283688 BJ283688
c 43	87	13.3	643	12	BJ254998	BJ254998 BJ254998
44	87	13.3	649	13	BQ801199	BQ801199 WHE2811_E
45	87	13.3	656	13	BU016293	BU016293 QHE12L01.

ALIGNMENTS

RESULT 1
U82114
LOCUS U82114 289 bp DNA linear GSS 16-FEB-2001
DEFINITION U82114 ordered cosmid library Mycobacterium leprae genomic clone cosmid L-373; contig 64, genomic survey sequence.
ACCESSION U82114
VERSION U82114.1 GI:3647212
KEYWORDS GSS.
SOURCE Mycobacterium leprae
ORGANISM Mycobacterium leprae
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
REFERENCE 1 (bases 1 to 289)

```
  AUTHORS   Silbaq,F.S., Cho,S.N., Cole,S.T. and Brennan,P.J.
  TITLE     Characterization of a 34-kilodalton protein of Mycobacterium leprae
            that is isologous to the immunodominant 34-kilodalton antigen of
            Mycobacterium paratuberculosis
  JOURNAL   Infect. Immun. 66 (11), 5576-5579 (1998)
  MEDLINE   99003183
   PUBMED   9784577
COMMENT     Contact: Silbaq FS
            Microbiology
            Colorado State University
            Fort Collins, CO 80523, USA
            Eiglmeier,K., Honore,N., Woods,S.A., Caudron,B. and Cole,S.T. Use
            of an ordered cosmid library to deduce the genomic organization of
            Mycobacterium leprae. Mol. Microbiol. 7 (2), 197-206 (1993)
            Class: unknown.
FEATURES             Location/Qualifiers
     source          1. .289
                     /organism="Mycobacterium leprae"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1769"
                     /clone="cosmid L-373; contig 64"
                     /clone_lib="ordered cosmid library"
ORIGIN

Alignment Scores:
Pred. No.:              1.19e-05        Length:         289
Score:                  145.50          Matches:        35
Percent Similarity:     58.82%          Conservative:   15
Best Local Similarity:  41.18%          Mismatches:     34
Query Match:            22.28%          Indels:         1
DB:                     29              Gaps:           1

US-09-684-215B-18 (1-128) x U82114 (1-289)

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsn---GlyAlaArgValGlnArg 59
              |||   ||||||:::  |||     :::  |||      |||||||:::|||
Db          7 GCGTCGCTGGGTGTACAGGTGGCCACCGACAAAGGCACCCCGGGCGCCAAAGTTATGGAC 66

Qy         60 ValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAla 79
              ||||||          ||||||:::     :::      |||   |||:::|||
Db         67 GTCGTGGCCGGTGGTGCGGCCGCGAATGCCGCCGTTCCCAAGGGGGTCGTTCTCACTAAG 126

Qy         80 ValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHis 99
              ||||||        |||:::||||||   |||:::      |||:::
Db        127 GTCGACGACCGCCTGATCAGTAGCGCCGACGCGTTGGTCGCTGCCGTGCGGTCCAAGGCA 186

Qy        100 ProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsn 119
              |||||||||   :::|||:::|||:::|||   :::||||||   :::|||||||
Db        187 CCCGGTGACAAAGTGTCGCTGACCTATCAGGATCAGTCTGGTAGCAGTCGCACGGTTCAG 246

Qy        120 ValThrLeuAlaGlu 124
              |||||||||   :::
Db        247 GTCACACTCGGCAAG 261

RESULT 2
BH770798/c
LOCUS       BH770798               1438 bp    DNA     linear   GSS 01-MAY-2002
DEFINITION  LLMGtag541 MG1363 Random Sequence Tag Library Lactococcus lactis
            subsp. cremoris genomic, genomic survey sequence.
ACCESSION   BH770798
VERSION     BH770798.1  GI:20373755
KEYWORDS    GSS.
SOURCE      Lactococcus lactis subsp. cremoris
  ORGANISM  Lactococcus lactis subsp. cremoris
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Lactococcus.
REFERENCE   1  (bases 1 to 1438)
  AUTHORS   Bolotin,A., Ehrlich,S.D. and Sorokin,A.
  TITLE     Studies of genomes of dairy bacteria Lactococcus lactis
  JOURNAL   Sci. Aliments (2002) In press
COMMENT     Contact: Sorokin A
            Genetique Microbienne
            INRA
            CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
            Tel: 33 1 34 65 25 16
            Fax: 33 1 34 65 25 21
            Email: sorokine@jouy.inra.fr
            best homologue in strain IL1403 is htrA (95%)
            Class: shotgun
            High quality sequence start: 30
            High quality sequence stop: 1408.
FEATURES             Location/Qualifiers
     source          1. .1438
                     /organism="Lactococcus lactis subsp. cremoris"
                     /mol_type="genomic DNA"
                     /strain="MG1363"
                     /sub_species="cremoris"
                     /db_xref="taxon:1359"
                     /clone_lib="MG1363 Random Sequence Tag Library"
                     /note="Vector: pSGMU2; Site_1: SmaI; Library of
                     chromosomal fragments of L.lactis strain MG1363 was
                     prepared by partial AluI digestion or by sonication."
ORIGIN

Alignment Scores:
Pred. No.:              0.0172          Length:         1438
Score:                  125.50          Matches:        40
Percent Similarity:     43.28%          Conservative:   18
Best Local Similarity:  29.85%          Mismatches:     45
Query Match:            19.22%          Indels:         31
DB:                     28              Gaps:           5

US-09-684-215B-18 (1-128) x BH770798 (1-1438)

Qy         13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla---------------- 27
              |||   |||||||||||||||             :::   |||
Db        479 GGTTTAGGTTTTGCTATCCCATCTAATGATGTGGTAAACATCATTAATAAACTTGAAACT 420

Qy         28 ---GlyGlnIleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGly 46
                 |||:::|||      |||                              ||||||:::
Db        419 GATGGTAAGATTTCACGTCCTGCC-----------------------TTAGGTATTCGT 384

Qy         47 ValValAsp-----------------AsnAsnGlyAsn---------------------- 53
              :::||||||               |||:::   :::
Db        383 ATGGTTGACCTATCTCAATTATCAACAAATGATAGTTCTCAACTGAAATTACCTAGCAGC 324

Qy         54 ---------GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGly 70
                      |||     |||     |||                ||||||:::   |||
Db        323 GTAACTGGTGGAGTGGTTGTCTACTCTGTTCAAGCGGGTCTTTCTGCTGCCACAGCTGGT 264

Qy         71 IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAla 90
              :::      |||||||||||||||   |||              :::   |||:::|||
Db        263 CTGAAAGCTGGCGATGTGATAACGAAGGTGGGAGATACCGCCGTTACTTCATCAACAGAC 204

Qy         91 MetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThr 110
              :::      ||||||      |||:::      |||   :::   ||||||:::
Db        203 TTACAAAGTGCTCTTTACTCACACAATATTAATGATACTGTGAAAGTCACCTACTAC--- 147

Qy        111 LysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGlu 124
              :::   |||  :::   |||  ||||||   |||::::::
Db        146 CGTGATGGTAAATCAGCCACAGCAAATGTCAAACTCTCTAAA 105

RESULT 3
BZ549048/c
LOCUS       BZ549048               959 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  pacs1-60_1644.s2 pacs1-60 Pseudomonas aeruginosa genomic clone
            pacs1-60_1644, genomic survey sequence.
ACCESSION   BZ549048
VERSION     BZ549048.1  GI:27152629
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 959)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
```

```
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .959
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="1-60"
                     /db_xref="taxon:287"
                     /clone="pacs1-60_1644"
                     /clone_lib="pacs1-60"
                     /note="clinical isolate 1-60 Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:                0.0171          Length:         959
Score:                    123.00          Matches:        41
Percent Similarity:       44.88%          Conservative:   16
Best Local Similarity:    32.28%          Mismatches:     46
Query Match:              18.84%          Indels:         24
DB:                       28              Gaps:           4

US-09-684-215B-18 (1-128) x BZ549048 (1-959)

Qy          4 SerAspAsnPheGlnLeuSerGlnGly------GlyGlnGlyPheAlaIleProIleGly 21
              |||       |||       |||   |||       |||       ||||||||||||||
Db        378 TCGCAGATCTTCACCCGTTCCGGCGGCTTCATGGGCCTGTCCTTCGCCATTCCGATCGAT 319

Qy         22 GlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThrAla 41
                ||:::   :::|||  |||:::|||                    |||     :::
Db        318 GTCGCGCTGAACGTCGCCGACCAGTTGAAG------------AAAGCCGGCAAGGTCAGT 271

Qy         42 PheLeuGlyLeuGlyValValAspAsnAsnGlyAsn------------------------ 53
                        ||||||||||||          |||
Db        270 CGCGGCTGGCTGGGCGTGGTGATCCAGGAAGTGAACAAGGATCTCGCCGAGTCCTACGGC 211

Qy         54 ---------------GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSer 68
                          ||||||   |||   :::::::|||          |||||||||
Db        210 CTCGACAAGCCGTCCGGCGCGCTGGTGGCGCAACTGGTGGAAGACGGCCCGGCGGCCAAG 151

Qy         69 LeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 88
                ||:::      ||||||||||||   :::::::::|||       ||||||   :::
Db        150 GGTGGCCTGCAGGTGGGCGATGTGATCATAAGCTTGAACGGCCAGTCGATCAACGAGTCC 91

Qy         89 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 108
                  :::         :::           |||||||||   |||::::::
Db         90 GCCGACCTGCCGCACCTGGTGGGCAACATGAAGCCGGGCGACAAGATCAACCTG------ 37

Qy        109 GlnThrLysSerGlyGlyThr 115
                       ||||||:::
Db         36 ---------GACGGGGGATCC 25

RESULT 4
CB679186/c
LOCUS       CB679186                726 bp    mRNA    linear   EST 09-APR-2003
DEFINITION  OSJNEf02F19.r OSJNEf Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEf02F19 3', mRNA sequence.
ACCESSION   CB679186
VERSION     CB679186.1  GI:29682911
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 02  row: P  column: 19
            Seq primer: gga aac agc tat gac cat g.
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEf02F19"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEf"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Uninfected Control"
ORIGIN

Alignment Scores:
Pred. No.:                0.0126          Length:         726
Score:                    122.50          Matches:        41
Percent Similarity:       42.86%          Conservative:   16
Best Local Similarity:    30.83%          Mismatches:     49
Query Match:              18.76%          Indels:         27
DB:                       14              Gaps:           4

US-09-684-215B-18 (1-128) x CB679186 (1-726)

Qy         10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
                :::       |||   ||||||||||||||||                :::   ||||||   |||
Db        579 ACATCTGCTGGTGTTGGTTTTGCCATCCCGTCATCAACTGTTCTGAAAATAGCTCCTCAG 520

Qy         30 IleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValValAsp 49
                ::::             :::       |||                |||||   |||
Db        519 TTA-----------ATTCAGTTTGGAAAAGTTCGCCGTGCTGGCTTGAATGTGGAGTTT 472

Qy         50 ---------------------------AsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
                                        |||          |||:::   :::   :::|||
Db        471 GCTCCAGATCCAATCGCATATCAGCTTAATGTTCGCACTGGATCTATAATATTGCAGGTT 412

Qy         61 ValGlySerAlaProAlaAlaSerLeuGly---------------------------- 70
                |||   :::   ||||||      |||
Db        411 CCTGGGGGCAGTGCTGCAGCCAAAGCTGGTCTTGTTCCTACCAGTAGGGGTTTTGCTGGT 352

Qy         71 ---IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThr 89
                  |||       ||||||||||||   ||||||||||||   ||||||       :::
Db        351 ACCATTGTTCTTGGTGATGTCATTGTTGCCGTGGACGGTAAACCTATCAAAGGCAAATCT 292

Qy         90 AlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGln 109
                ::::::      |||:::   :::      ||||||   :::|||:::|||   |||
Db        291 GATCTATCGAGGGTTCTGGATGCTTATGGCGTTGGAGACAAGGTGAGCTTGACAATCCAA 232

Qy        110 ThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 122
                       |||       :::       |||||||||
Db        231 AGA---------GGCGCTGAATCTCTGGAGGTAACCTTG 202
```

```
RESULT 5
CB925602
LOCUS       CB925602                590 bp    mRNA    linear   EST 28-APR-2003
DEFINITION  ABA1_22_F06.b1_A012 Abscisic acid-treated seedlings Sorghum bicolor
            cDNA clone ABA1_22_F06_A012 3', mRNA sequence.
ACCESSION   CB925602
VERSION     CB925602.1  GI:30161873
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Cordonnier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S.,
            Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Buchanan,C.D.,
            Eastman,A. and Pratt,L.H.
  TITLE     An EST database from Sorghum: ABA1-treated seedlings
  JOURNAL   Unpublished (2003)
COMMENT     Other_ESTs: ABA1_22_F06.g1_A012
            Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
            the Human Genome Center, University of Tokyo Institute of Medical
            Science; plant material and RNA prepared at Texas A & M University;
            sequencing done in the Laboratory for Genomics and Bioinformatics,
            University of Georgia. Sequence ends have been trimmed to exclude
            vector and regions below Phred quality 16. Three-prime sequences
            are presented as their reverse complement and have been trimmed to
            exclude polyA.
            Seq primer: Sug3 (CGACCTGCAGCTCGAGCACA)
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /cultivar="IS3620C"
                     /db_xref="taxon:4558"
                     /clone="ABA1_22_F06_A012"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="Abscisic acid-treated seedlings"
                     /note="Vector: pME18S-FL3; Site_1: XhoI; Site_2: XhoI; The
                     library was prepared from polyA+ RNA from seedlings grown
                     in hydroponic culture. After 12 days, medium was
                     supplemented with 1 mM abscisic acid (ABA), while leaves
                     were misted with a solution of 1 mM ABA. Roots and leaves
                     were harvested after 3, 6, 12, and 24 hr and material from
                     all time points was combined prior to RNA isolation.
                     Double-stranded cDNA was cloned unidirectionally into
                     different DraIII sites of the pME18S-FL3 vector (5-prime
                     DraIII site is CACTGTGTG, 3-prime DraIII site is
                     CACCATGTG). XhoI excises the cDNA insert."
ORIGIN

Alignment Scores:
Pred. No.:              0.0191        Length:         590
Score:                  119.50        Matches:        38
Percent Similarity:     43.41%        Conservative:   18
Best Local Similarity:  29.46%        Mismatches:     40
Query Match:            18.30%        Indels:         33
DB:                     14            Gaps:           5

US-09-684-215B-18 (1-128) x CB925602 (1-590)

Qy         15 GlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThr 34
              ||||||||||||||            :::   ||||||   ||:::
Db          6 GGCTTTGCTATTCCATCATCAACTGTACTTAAAATCGCTCCTCAGTTA----------- 53

Qy         35 ValHisIleGlyProThrAlaPheLeuGlyLeuGlyValValAsp--------------- 49
              :::     |||             ||||||     ||||||
Db         54 ATTCAGTTCGGAAAAGTTCGTCGTGCTGGCTTGAAT---GTGGACTTCGCTCCGGATCCA 110

Qy         50 --------------AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAla 64
                           |||       |||||||||   :::   :::|||   |||   :::
Db        111 ATTGCATATCAGCTTAATGTTCGCAACGGAGCTCTTATACTTAAGGTACCTGGGGGCAGT 170

Qy         65 ProAlaAlaSerLeuGlyIleSer---------------------------------Thr 73
              ||||||      |||::::::
Db        171 GCTGCAGCCAAAGCAGGTCTTGCTCCAACCGGCAGGGGTTTTGCTGGTAATATTGTTCTG 230

Qy         74 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 93
              ||||||:::|||   ||||||||||||   |||:::         :::   :::
Db        231 GGTGATATCATCGTTGCAGTGGACGGCAAACCTGTTAAGGGCAAATCTGACCTGCTGAGG 290

Qy         94 AlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGly 113
              |||:::   :::      ||||||   :::::::::|||
Db        291 GTTCTGGATGACTATGGCGTCGGAGATCAGGTGACCTTGACA----------------- 332

Qy        114 GlyThrArgThrGlyAsnValThrLeu 122
                 |||   |||:::   ||||||
Db        333 ---ATCCGGCGAGGCTCAGAAACCCTT 356

RESULT 6
CA228900/c
LOCUS       CA228900                807 bp    mRNA    linear   EST 25-SEP-2003
DEFINITION  SCQGFL3058G06.g Saccharum officinarum FL3 Saccharum officinarum
            cDNA clone SCQGFL3058G06 5', mRNA sequence.
ACCESSION   CA228900
VERSION     CA228900.1  GI:35290878
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE   1  (bases 1 to 807)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 058  row: G  column: 06
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .807
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCQGFL3058G06"
                     /lab_host="DH10B"
                     /clone_lib="Saccharum officinarum FL3"
                     /note="Organ: Base of developing inflorescence (5cm-long);
                     Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [Base of
                     developing  inflorescence (5cm-long)]. cDNA was prepared
                     from polyA+ mRNA using  SuperScript Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were  fractionated
                     in a sepharose CL-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details  of each source of RNA and library
                     construction can be obtained at
                     http://sucest.lad.ic.unicamp.br/public"
```

```
ORIGIN

Alignment Scores:
Pred. No.:                0.0499          Length:         807
Score:                    117.50          Matches:        39
Percent Similarity:       42.54%          Conservative:   18
Best Local Similarity:    29.10%          Mismatches:     44
Query Match:              17.99%          Indels:         33
DB:                       14              Gaps:           5

US-09-684-215B-18 (1-128) x CA228900 (1-807)

Qy          10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
               :::      |||   ||||||||||||              :::   ||||||   |||
Db         569 ACATCTGCTGGTGTTGGCTTCGCTATCTCATCATCAACTGTGCTTAAAATCGCTCCTCAG 510

Qy          30 IleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValValAsp 49
               :::           :::     |||                ||||||     ||||||
Db         509 TTA-----------ATTCAGTTTGGAAAAGTTCGTCGTGCTGGCTTGAAT---GTGGAC 465

Qy          50 ------------------------------AsnAsnGlyAsnGlyAlaArgValGlnArg 59
                                           |||      |||||||||   :::   :::
Db         464 TTTGCTCCAGATCCAATTGCGTATCAGCTTAATGTTCGCAACGGAGCTCTTATACTTAAG 405

Qy          60 ValValGlySerAlaProAlaAlaSerLeuGlyIleSer---------------------- 72
               |||   |||   :::   ||||||      |||::::::
Db         404 GTACCTGGGGGCAGTGCTGCAGCCAAGGCAGGCCTTGCTCCGACCGGCAGGGGTTTCGCT 345

Qy          73 ------------ThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 88
                           ||||||||||||   ||||||||||||   |||:::
Db         344 GGTAACATTGTTCTTGGTGATGTCATCGTTGCAGTGGATGGCAAACCTGTTAAGGGCAAA 285

Qy          89 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 108
               :::   :::      |||:::   :::      ||||||   :::::::::|||
Db         284 TCTGACCTGCTGAGGGTTCTGGATGACTATGGCGTCGGGGATCAGGTGACCTTGACA--- 228

Qy         109 GlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 122
                        |||   |||:::   ||||||
Db         227 -----------------ATCCGGCGAGGCTCTGAAACCCTC 204

RESULT 7
BQ514888/c
LOCUS       BQ514888                758 bp    mRNA    linear   EST 07-MAR-2003
DEFINITION  EST622303 Generation of a set of potato cDNA clones for microarray
            analyses mixed potato tissues Solanum tuberosum cDNA clone STMIO81
            3' end, mRNA sequence.
ACCESSION   BQ514888
VERSION     BQ514888.1  GI:21373757
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
            Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
            Karamycheva,S.A.
  TITLE     Generation of a set of potato cDNA clones for microarray analyses
  JOURNAL   Unpublished (2002)
COMMENT     Other_ESTs: EST622302
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .758
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec or Binjte"
                     /db_xref="taxon:4113"
                     /clone="STMIO81"
                     /tissue_type="mixed tissues"
                     /lab_host="SOLR"
                     /clone_lib="Generation of a set of potato cDNA clones for
                     microarray analyses mixed potato tissues"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Combination of untreated and Phytophthora
                     infestans-treated libraries of stolons, leaves, leaflets,
                     axillary buds of stem explants, petioles, germinating
                     eyes, tubers, or roots."
ORIGIN

Alignment Scores:
Pred. No.:                0.106           Length:         758
Score:                    114.00          Matches:        34
Percent Similarity:       42.61%          Conservative:   15
Best Local Similarity:    29.57%          Mismatches:     42
Query Match:              17.46%          Indels:         24
DB:                       13              Gaps:           3

US-09-684-215B-18 (1-128) x BQ514888 (1-758)

Qy          10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
               :::      |||   ||||||||||||||||              :::   :::      |||
Db         592 ACATCAGCAGGTGTTGGATTTGCAATCCCTTCTTCAACTGTGTTGAGAGTTGTGCCCCAG 533

Qy          30 IleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValValAsp 49
               :::           :::     |||                ||||||   :::
Db         532 TTG-----------ATCCAATCTGGAAAAGTTCTTCGTGCTGGTTTGAATATTGAAATC 485

Qy          50 ---------------------------AsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
                                          |||      |||||||||   |||      |||
Db         484 GCTCCAGACCTGATTGCCAACCAACTTAATGTTCGAAATGGAGCACTGGTTTTGCTGGTA 425

Qy          61 ValGlySerAlaProAlaAlaSerLeuGly----------------------------- 70
               |||::::::|||||||||      |||
Db         424 CCTGGAAATAGTCCTGCAGCGAAAGCGGGACTTCTTCCTACTACCAGGGGTTTTGCAGGA 365

Qy          71 ---IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThr 89
                  |||      ||||||:::|||   |||||||||      |||:::   |||
Db         364 AATATAGTGCTTGGCGATATTATTGAAGCAGTGGATGACAAACCTGTTAGGAGTAAAGCA 305

Qy          90 AlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIle 104
               :::      ||||||:::   ::::::   ||||||   :::
Db         304 GAGTTGTATAAAGCCCTGGATAACTATAACATAGGTGATGAAGTT 260

RESULT 8
AZ934428
LOCUS       AZ934428                603 bp    DNA     linear   GSS 24-APR-2001
DEFINITION  BJ__Ba0002I08r B. japonicum BAC library Bradyrhizobium japonicum
            genomic, genomic survey sequence.
ACCESSION   AZ934428
VERSION     AZ934428.1  GI:13776488
KEYWORDS    GSS.
SOURCE      Bradyrhizobium japonicum
  ORGANISM  Bradyrhizobium japonicum
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE   1  (bases 1 to 603)
  AUTHORS   Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A.,
            Goicoechea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.
  TITLE     A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
            genome
  JOURNAL   Genome Res. 11 (8), 1434-1440 (2001)
  MEDLINE   21376150
   PUBMED   11483585
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
```

```
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Class: BAC ends
            High quality sequence stop: 553.
FEATURES              Location/Qualifiers
     source           1. .603
                      /organism="Bradyrhizobium japonicum"
                      /mol_type="genomic DNA"
                      /strain="USDA110"
                      /db_xref="taxon:375"
                      /lab_host="E. coli"
                      /clone_lib="B. japonicum BAC library"
                      /note="Vector: pIndigo536; Site_1: HindIII"
ORIGIN

Alignment Scores:
Pred. No.:                0.522          Length:         603
Score:                    106.00         Matches:        35
Percent Similarity:       41.80%         Conservative:   16
Best Local Similarity:    28.69%         Mismatches:     55
Query Match:              16.23%         Indels:         16
DB:                       28             Gaps:           4

US-09-684-215B-18 (1-128) x AZ934428 (1-603)

Qy          13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleLys--- 31
               |||     |||:::||||||                :::     |||:::|||
Db         245 GGCATCGCGTTCTCGATCCCGGCCAACACCGTGAAGACGGTTGTTGCCCAGCTCAAGGAC 304

Qy          32 --------------------LeuProThrValHisIleGlyProThrAlaPheLeuGly 44
                                   |||   |||   |||
Db         305 AAGGGTTCGGTCAGCCGCGGCTGGATCGGCGTGCAGATTCAGCCGGTGACG--------- 355

Qy          45 LeuGlyValValAspAsnAsnGly----------AsnGlyAlaArgValGlnArgVal 60
                   :::   |||:::   |||              ||||||   |||
Db         356 TCGGATATCGCCGACAGCCTCGGCATGAAGAAGGCCGAAGGGGCGCTGGTGGCGGAGCCG 415

Qy          61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
                   :::   |||||||||     ||||||   :::|||||||||||||||:::|||
Db         416 CAGGCGAACGGTCCGGCGGCGAAGGCCGGCATCGAGTCCGGCGACGTGATCACCTCGGTC 475

Qy          81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
               :::|||      :::     |||     :::|||     :::   |||       |||
Db         476 AACGGCGAATCCGTCAAGGACGCCCGCGAGCTCGCCCGCACCATCGGCGGCATGGCGCCC 535

Qy         101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
                   ::::::   :::          |||   |||                |||:::
Db         536 CGTGCGATCGTGAAGCTTAACGTGCTGCACAAG---GGCCAGGACAATGTCGTGAACCTC 592

Qy         121 ThrLeu 122
                  ||||||
Db         593 ACCCTC 598

RESULT 9
AJ558965
LOCUS       AJ558965                765 bp    mRNA    linear   EST 12-JUN-2003
DEFINITION  AJ558965 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
            018_1_09_a08, mRNA sequence.
ACCESSION   AJ558965
VERSION     AJ558965.1  GI:31661537
KEYWORDS    EST.
SOURCE      Antirrhinum majus (snapdragon)
  ORGANISM  Antirrhinum majus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Lamiales; Antirrhinaceae; Antirrhineae;
            Antirrhinum.
REFERENCE   1  (bases 1 to 765)
  AUTHORS   Zachgo,S., Stueber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.
  TITLE     Antirrhinum EST collection
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Schwarz-Sommer Z
            Molekulare Pflanzengenetik
            MPI fuer Zuechtungsforschung
            Carl-von-Linne Weg 10, D-50829, Germany.
FEATURES              Location/Qualifiers
     source           1. .765
                      /organism="Antirrhinum majus"
                      /mol_type="mRNA"
                      /db_xref="taxon:4151"
                      /clone="018_1_09_a08"
                      /tissue_type="whole plant"
                      /clone_lib="Antirrhinum majus whole plant"
ORIGIN

Alignment Scores:
Pred. No.:                1.21           Length:         765
Score:                    104.00         Matches:        34
Percent Similarity:       42.24%         Conservative:   15
Best Local Similarity:    29.31%         Mismatches:     41
Query Match:              15.93%         Indels:         26
DB:                       9              Gaps:           4

US-09-684-215B-18 (1-128) x AJ558965 (1-765)

Qy          10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
               :::     |||   |||||||||||||||              :::     |||     |||
Db         311 ACCTCGGCGGGTGTAGGATTTGCGATCCCTTCATCGACAGTGGTAAAGATAGTGCCTCAG 370

Qy          30 IleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValValAsp 49
               :::             ::::::   |||                 ||||||     ||||||
Db         371 TTG------------ATCAATTATGGAAAAGTTGTTCGAGCTGGTTTAAAT---GTTGAC 415

Qy          50 ----------------------------AsnAsnGlyAsnGlyAlaArgValGlnArg 59
                                           |||      |||||||||   :::   :::
Db         416 ATAGCTCCTGATCTTATTGCAAATCAACTCAATGTTCGAAATGGAGCTCTAATCTTGCAG 475

Qy          60 ValValGlySerAlaProAlaAlaSerLeuGly-------------------------- 70
               |||   |||::::::   ||||||      |||
Db         476 GTCCCTGGAAATAGTTCAGCAGCTAAGGCTGGGCTTGTACCTACAACACGGGGTTTCGCT 535

Qy          71 ------IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 88
                     |||     ||||||:::|||   |||||||||      |||:::   |||
Db         536 GGTAAAATCGTACTTGGGGATATAATTTTGGCCGTAGATGACAAACCTGTCAGGAGTAGG 595

Qy          89 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIle 104
                   :::             |||:::   :::       |||||   :::
Db         596 GCCGAGCTATACAAAACACTGGATGAGTATTCCATCGGAGACAAAGTG 643

RESULT 10
BZ561390/c
LOCUS       BZ561390                1033 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  pacs2-164_3239.y3 pacs2-164 Pseudomonas aeruginosa genomic clone
            pacs2-164_3239, genomic survey sequence.
ACCESSION   BZ561390
VERSION     BZ561390.1  GI:27181349
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 1033)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES              Location/Qualifiers
```

```
     source          1. .1033
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="2-164"
                     /db_xref="taxon:287"
                     /clone="pacs2-164_3239"
                     /clone_lib="pacs2-164"
                     /note="clinical isolate 2-164 Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:              4.46          Length:         1033
Score:                  100.50        Matches:        42
Percent Similarity:     42.75%        Conservative:   17
Best Local Similarity:  30.43%        Mismatches:     54
Query Match:            15.39%        Indels:         25
DB:                     28            Gaps:           4

US-09-684-215B-18 (1-128) x BZ561390 (1-1033)

Qy          4 SerAspAsnPheGlnLeuSerGlnGlyGlyGln------GlyPheAlaIleProIleGly 21
              |||      |||      |||   |||             |||   |||||||||
Db        713 TCCCAGTTTTTCCACCGGTCCGCCGGCTCAATGGCCCTGTCTTTCGGCATTCCGATCGAT 654

Qy         22 GlnAlaMetAlaIle-AlaGlyGlnIleLys-LeuProThrValHisIleGlyProThrA 41
                ||:::   ::: |||   |||:::|||   |||             |||       :
Db        653 GTCGCGCTGACCGTTCGCCGACCAGTTGAAGAAAGCCC-----------GGCAAGGTCA 606

Qy         41 laPheLeuGlyLeuGlyValValAspAsnAsnGlyAsn---------------------- 53
              ::        |||||||||||||          |||
Db        605 GTCGCGGCTGGCTGGGCGTGGTGATCCAGGAAGTGAACAAGGATCTCGCCGAGTCCTTCG 546

Qy         54 -----------------GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaS 68
                              ||||||   |||   :::::::|||         |||||||||
Db        545 GCCTCGACAAGCCGTCCGGCGCGCTGGTGGCGCAACTGGTGGAAGACGGCCCGGCGGCCA 486

Qy         68 erLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerA 88
                ||:::      ||||||||||||   :::::::::|||      ||||||   :
Db        485 AGGGTGGCCTGCAGGTGGGCGATGTGATCCTCAGCCTGAACGGCCAGTCGATCAACGAGT 426

Qy         88 laThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal---- 106
              ::    :::      :::          |||||||||   |||::::::
Db        425 CCGCCGACCTGCCGCACCTGGTGGGCAACATGAAGCCGGGCGACAAGATCAACCTGGACG 366

Qy        107 --------ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
                      |||   ||||||          |||::::::
Db        365 TGATTCGCAACGGCCAGCGCAAGTCCCTGAGCATGGCGGTAGGCAGCCTT 316

RESULT 11
BZ549047/c
LOCUS       BZ549047                947 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  pacs1-60_1644.s1 pacs1-60 Pseudomonas aeruginosa genomic clone
            pacs1-60_1644, genomic survey sequence.
ACCESSION   BZ549047
VERSION     BZ549047.1  GI:27152628
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 947)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .947
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="1-60"
                     /db_xref="taxon:287"
                     /clone="pacs1-60_1644"
                     /clone_lib="pacs1-60"
                     /note="clinical isolate 1-60 Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:              4.42          Length:         947
Score:                  100.00        Matches:        38
Percent Similarity:     45.61%        Conservative:   14
Best Local Similarity:  33.33%        Mismatches:     43
Query Match:            15.31%        Indels:         20
DB:                     28            Gaps:           3

US-09-684-215B-18 (1-128) x BZ549047 (1-947)

Qy          4 SerAspAsnPheGlnLeuSerGlnGly------GlyGlnGlyPheAlaIleProIleGly 21
              |||      |||      |||   |||      |||       ||||||||||||||||
Db        374 TCGCAGATCTTCACCCGTTCCGGCGGCTTCATGGGCCTGTCCTTCGCCATTCCGATCGAT 315

Qy         22 GlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThrAla 41
                ||||||   :::|||   |||:::|||                    |||       :::
Db        314 GTCGCGATGAACGTCGCCGACCAGTTGAAG------------AAAGCCGGCAAGGTCAGT 267

Qy         42 PheLeuGlyLeuGlyValValAsp---------------------------------- 49
                        |||||||||   |||
Db        266 CGCGGCTGGCTGGGCGTGGT-GATCCAGGAAGTGAACAACGATCTCGCCGAGTCTTCGGC 208

Qy         50 ---AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSer 68
                  :::      :::|||||||   |||   ::::::|||         |||||||||
Db        207 CTCGACAAGCCGTCCGGCGCGCTGGTGGCGCAACTGGTGGAAGACGGCCCGGCGGCGAAG 148

Qy         69 LeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 88
                ||:::      ||||||||||||   :::::::::|||      ||||||   :::
Db        147 GGTGGCCTGCAGGTGGGCGATGTGATCAACAGCCTGAACGGCCAGTCGATCAACGAGTCC 88

Qy         89 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAsp 102
                  :::      :::          |||||||||
Db         87 GCCGACCTGCCGCACCTGGTGTGGAACATGAAGCCGGGCGAC 46

RESULT 12
AQ989479
LOCUS       AQ989479                726 bp    DNA     linear   GSS 14-AUG-2000
DEFINITION  Rfc00025 Photorhabdus luminescens strain W14 M13 library
            Photorhabdus luminescens genomic clone PLG00025, genomic survey
            sequence.
ACCESSION   AQ989479
VERSION     AQ989479.1  GI:9648073
KEYWORDS    GSS.
SOURCE      Photorhabdus luminescens
  ORGANISM  Photorhabdus luminescens
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Photorhabdus.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   ffrench-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,
            Daborn,P.J., Bowen,D. and Blattner,F.R.
  TITLE     A genomic sample sequence of the entomopathogenic bacterium
            Photorhabdus luminescens W14: potential implications for virulence
  JOURNAL   Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
  MEDLINE   20378633
   PUBMED   10919786
COMMENT     Contact: ffrench-Constant RH
            Department of Biology and Biochemistry
            University of Bath
            South Building, Bath BA2 7AY, UK
```

Tel: (44) 1225 826621
Fax: (44) 1225 826779
Email: bssrfc@bath.ac.uk
This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic Acids Res.
Seq primer: M13 Forward
Class: shotgun.

FEATURES Location/Qualifiers
source 1. .726
/organism="Photorhabdus luminescens"
/mol_type="genomic DNA"
/strain="W14"
/db_xref="taxon:29488"
/clone="PLG00025"
/dev_stage="primary phase variant"
/clone_lib="Photorhabdus luminescens strain W14 M13 library"
/note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."

ORIGIN

Alignment Scores:

Pred. No.:	3.77	Length:	726
Score:	99.00	Matches:	38
Percent Similarity:	45.83%	Conservative:	17
Best Local Similarity:	31.67%	Mismatches:	43
Query Match:	15.16%	Indels:	23
DB:	28	Gaps:	5

US-09-684-215B-18 (1-128) x AQ989479 (1-726)

```
Qy          6 AsnPheGlnLeuSerGlnGlyGlyGln----------------GlyPheAlaIleProIle 20
              :::|||      |||:::   |||:::                |||||||||||||||
Db        179 TCATTCGATAAAAGTGAAAATGGTGAAACGCCAGAAGGGTTGGGGTTCGCTATTCCG--- 235

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
                        :::|||   :::|||          :::      |||
Db        236 ---------ACTGAACTCGCTACTAAGATTATGCAGAAACTTATCCGTGACGGACGAGTT 286

Qy         41 ------AlaPheLeuGlyLeu-----------------GlyValValAspAsnAsnGly 52
                  |||:::|||:::                    :::     :::||||||
Db        287 ATCCGGGGTTTTATCGGCATAACATCGCAAGAGCTGCCACATATTCGCTCTTCTAACGGC 346

Qy         53 Asn---------------GlyAlaArgValGlnArgValValGlySerAlaProAlaAla 67
              |||              |||  ||||||   :::|||    :::   ||||||
Db        347 AATATTAACCAGCTTCAAGGGCTACGAGTATTTCAGGTAACTACAAACGGGCCTGCGCAA 406

Qy         68 SerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSer 87
                 ||||||      ||||||:::||||||:::|||:::      |||      |||
Db        407 AAAGCCGGTATTAAAGTGGGTGATATCATTACCAGCGTCAATAACAAACCCGCTATTTCT 466

Qy         88 AlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThr 107
              |||           |||  :::     :::  ||||||  |||     ||||||
Db        467 GCGGGTGAAACAATGGATCAAGTGGCTGA-AATACCCCCGGCAGCGTGGGTNCTGTCACT 525
```

RESULT 13
BU557763/c
LOCUS BU557763 1093 bp mRNA linear EST 16-SEP-2002
DEFINITION AGENCOURT_10214312 NIH_MGC_107 Homo sapiens cDNA clone IMAGE:6585799 5', mRNA sequence.
ACCESSION BU557763
VERSION BU557763.1 GI:22908059
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1093)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2798 row: f column: 07
High quality sequence stop: 299.

FEATURES Location/Qualifiers
source 1. .1093
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6585799"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/note="Organ: breast; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	8.91	Length:	1093
Score:	98.00	Matches:	44
Percent Similarity:	37.04%	Conservative:	16
Best Local Similarity:	27.16%	Mismatches:	41
Query Match:	15.01%	Indels:	62
DB:	13	Gaps:	8

US-09-684-215B-18 (1-128) x BU557763 (1-1093)

```
Qy         26 IleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThrAla---------- 41
              :::||||||::::::      ||||||      |||||||||
Db        750 ATGGCGGGGCGGCTAAGCTATCCCACGGCGCCCCGCGGCCCCACGGTGCAGAGCCCCCAG 691

Qy         42 -------------------------------------PheLeuGlyLeuGlyValVal 48
                                                  |||   |||:::
Db        690 GGCGGCCGGGGCACGCCTCGGGCCAAGGGGGGCACCTGGAGTAGGGGGTACGGGATTTGC 631

Qy         49 AspAsnAsnGlyAsnGlyAlaArgValGln---------------------------- 58
                 ||||||:::||||||||:::|||
Db        630 ---GGCAATGGCTCAGGCGCTCGCCTACAACCCGTAAAACCAGTCTCCGGTGGAGGATCT 574

Qy         59 ------ArgValValGly-----------SerAlaProAlaAlaSerLeuGlyIleSer 72
                   |||:::|||                |||   |||      |||::::::
Db        573 CTCGGGGAGGTAATTGGGTGCAGGGACGCATGCCAACCAGGGGCCGGGCGCGGAGTGGCG 514

Qy         73 ThrGlyAspValIleThrAlaValAsp-----------GlyAlaProIleAsnSerAla 88
              |||||||||   :::      |||:::           |||   |||
Db        513 ACGGGGGACGGCCTACACCTGGTAGAGTGGCACCGGGGTGGGGGACCC---CGTGTCCAG 457

Qy         89 ThrAlaMetAlaAspAla----------------------------------LeuAsn 96
              |||:::|||   ||||||                                  |||:::
Db        456 ACGTCCATGCCCGATGCCAGGGTGGCGCCCAGAGGGATAGAGGGACAGCACGGCCTGGAT 397

Qy         97 GlyHis----------------------------HisProGlyAspValIleSerVal 106
              ||||||                          |||   |||         |||:::
Db        396 GGCCACGTACATGACCGGGGGGTTGAAGGTGTCAAACATAATGGGAGTCA-TCTTCTCTC 338

Qy        107 ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGluGlyPro 126
                |||         |||||||||   |||:::      :::            ||||||
Db        337 TGTTGGCCTTGGGGTTCAGGGGGACCTCGGTCAGCAGCACTGGGTGCTCCTCCGGGGCCA 278

Qy        127 ProAla 128
```

```
Db        277 CGCGCA 272

RESULT 14
CA148171
LOCUS       CA148171                444 bp    mRNA    linear   EST 24-SEP-2003
DEFINITION  SCEZRZ1017F10.g RZ1 Saccharum officinarum cDNA clone SCEZRZ1017F10
            5', mRNA sequence.
ACCESSION   CA148171
VERSION     CA148171.1  GI:35048941
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 017  row: F  column: 10
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCEZRZ1017F10"
                     /lab_host="DH10B"
                     /clone_lib="RZ1"
                     /note="Organ: Shoot-root transition zone from young plants
                     (large insert library); Vector: pSport1; Site_1: SalI;
                     Site_2: NotI; An unidirectional cDNA library generated
                     from [Shoot-root transition zone  from young plants (large
                     insert library)]. cDNA was prepared from polyA+  mRNA
                     using SuperScript Plasmid System Kit (Invitrogen). The
                     double-strand  cDNAs were fractionated in a sepharose
                     CL-2B 40cm-columns and fragments  sizing between 0.8 and
                     1.5 Kb were directionally cloned into the vector.  Details
                     of each source of RNA and library construction can be
                     obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

Alignment Scores:
Pred. No.:              2.91          Length:        444
Score:                  97.00         Matches:       26
Percent Similarity:     47.62%        Conservative:  14
Best Local Similarity:  30.95%        Mismatches:    26
Query Match:            14.85%        Indels:        18
DB:                     13            Gaps:          2

US-09-684-215B-18 (1-128) x CA148171 (1-444)

Qy          50 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 69
               |||       |||||||||  :::   :::|||   |||   :::   ||||||
Db          32 AATGTTCGCAACGGAGCTCTTATACTTAAGGTACCTGGGGGCAGTGCTGCAGCCAAGGCA 91

Qy          70 GlyIleSer--------------------------------ThrGlyAspValIleThr 78
               |||::::::                               ||||||||||||
Db          92 GGCCTTGCTCCGACCGGCAGGGGTTTCGCTGGTAACATTGTTCTTGGTGATGTCATCGTT 151

Qy          79 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 98
               ||||||||||||   |||:::         :::   :::         |||:::   :::
```

```
Db        152 GCAGTGGATGGCAAACCTGTTAAGGGCAAATCAGACCTGCTGAGGGTTCTGGATGACTAT 211

Qy         99 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 118
              ||||||   :::::::::|||                          |||   |||
Db        212 GGCGTCGGGGATCAGGTGACCTTGACA--------------------ATCCGGCGAGGC 250

Qy        119 AsnValThrLeu 122
              :::   ||||||
Db        251 TCTGAAACCCTC 262

RESULT 15
BZ577987
LOCUS       BZ577987                790 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  msh2_5667.x1 msh Pseudomonas aeruginosa genomic clone msh2_5667,
            genomic survey sequence.
ACCESSION   BZ577987
VERSION     BZ577987.1  GI:27213048
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 790)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .790
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="MSH"
                     /db_xref="taxon:287"
                     /clone="msh2_5667"
                     /clone_lib="msh"
                     /note="Environmental isolate. Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:              7.85          Length:        790
Score:                  96.50         Matches:       37
Percent Similarity:     42.40%        Conservative:  16
Best Local Similarity:  29.60%        Mismatches:    47
Query Match:            14.78%        Indels:        25
DB:                     28            Gaps:          6

US-09-684-215B-18 (1-128) x BZ577987 (1-790)

Qy          9 LeuSerGlnGlyGlyGlnGlyPheAlaIleProIleGly--------------GlnAla 23
              |||:::      ||||||      |||   |||   |||                |||
Db        110 CTTGCCGCTCCAGGGCAGCACCACGCGCACCCAGCAGGAGGAGGTCTCGTCGTTCTTGCC 169

Qy         24 MetAlaIleAlaGlyGlnIleLysLeu---ProThrValHisIleGlyProThrAlaPhe 42
              :::|||:::   |||::::::   |||   |||   |||   :::|||||
Db        170 GTCGCGGTCCCAGGGGAACTGGATCTTGATCCGGCCGTACTGGTCGGTCCAGATCTCCTC 229

Qy         43 LeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGly 62
              ||||||            ::::::      |||       |||          :::|||
Db        230 GCCGGCCTTGCC---------CACCACCAGGGCGGTCTGGGTGTGCATGCGCGGCTTGGG 280

Qy         63 SerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGly 82
              |||      :::::::|||          |||||   :::   ||||||
Db        281 TGTGACCCGCGCGGGGCGGTAGGCGGTGGCCTTGGGGATCGCCTCGAAGCGGTT------ 334
```

```
Qy          83 AlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAsp 102
               |||   :::      |||      :::   |||        |||   ||||||   |||
Db         335 GCGGTAGTGCGCGTGGCTGGCGTCGTGGGCGAC------CGAGGTCAGCACCCAGTCGAT 388

Qy         103 ValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 122
               :::      :::                                              |||
Db         389 ATTCAGCGAATCGTC----------------------------------GTCATGCCC 412

Qy         123 AlaGluGlyProPro 127
               |||||||||   |||
Db         413 GCTGAGGGTGAACCA 427

RESULT 16
AW285510
LOCUS       AW285510                 546 bp    mRNA    linear   EST 19-JUL-2000
DEFINITION  LG1_241_E05.g1_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   AW285510
VERSION     AW285510.1  GI:6675354
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 546)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
  TITLE     An EST database from Sorghum: light-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: T7
            High quality sequence start: 126
            High quality sequence stop: 546
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .546
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /db_xref="taxon:4558"
                     /clone_lib="Light Grown 1 (LG1)"
                     /note="Organ: 10- to 14-day-old light-grown (greenhouse)
                     seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2:
                     EcoRI; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
ORIGIN

Alignment Scores:
Pred. No.:              5.07          Length:         546
Score:                  96.00         Matches:        25
Percent Similarity:     47.62%        Conservative:   15
Best Local Similarity:  29.76%        Mismatches:     26
Query Match:            14.70%        Indels:         18
DB:                     10            Gaps:           2

US-09-684-215B-18 (1-128) x AW285510 (1-546)

Qy          50 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 69
               |||      |||||||||   :::   :::|||   |||   :::   ||||||
Db          56 AATGTTCGCAACGGAGCTCTTATACTTAAGGTACCTGGGGGCAGTGCTGCAGCCAAAGCA 115

Qy          70 GlyIleSer---------------------------------ThrGlyAspValIleThr 78
               |||::::::                                 ||||||:::|||
Db         116 GGTCTTGCTCCAACCGGCAGGGGTTTTGCTGGTAATATTGTTCTGGGTGATATCATCGTT 175

Qy          79 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 98
               ||||||||||||   |||:::         :::   :::         |||:::   :::
Db         176 GCAGTGGACGGCAAACCTGTTAAGGGCAAATCTGACCTGCTGAGGGTTCTGGATGACTAT 235

Qy          99 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 118
               ||||||   :::::::::|||                              |||   |||
Db         236 GGCGTCGGAGATCAGGTGACCTTGACA--------------------ATCCGGCGAGGC 274

Qy         119 AsnValThrLeu 122
               :::   ||||||
Db         275 TCAGAAACCCTT 286

RESULT 17
AW285527
LOCUS       AW285527                 551 bp    mRNA    linear   EST 19-JUL-2000
DEFINITION  LG1_241_G05.g1_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   AW285527
VERSION     AW285527.1  GI:6675371
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 551)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
  TITLE     An EST database from Sorghum: light-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: T7
            High quality sequence start: 16
            High quality sequence stop: 552
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .551
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /db_xref="taxon:4558"
                     /clone_lib="Light Grown 1 (LG1)"
                     /note="Organ: 10- to 14-day-old light-grown (greenhouse)
                     seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2:
                     EcoRI; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
ORIGIN

Alignment Scores:
Pred. No.:              5.14          Length:         551
Score:                  96.00         Matches:        25
Percent Similarity:     47.62%        Conservative:   15
Best Local Similarity:  29.76%        Mismatches:     26
Query Match:            14.70%        Indels:         18
DB:                     10            Gaps:           2

US-09-684-215B-18 (1-128) x AW285527 (1-551)

Qy          50 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 69
               |||      |||||||||   :::   :::|||   |||   :::   ||||||
Db          61 AATGTTCGCAACGGAGCTCTTATACTTAAGGTACCTGGGGGCAGTGCTGCAGCCAAAGCA 120

Qy          70 GlyIleSer---------------------------------ThrGlyAspValIleThr 78
```

```
            |||::::::                                        ||||||:::|||
Db        121 GGTCTTGCTCCAACCGGCAGGGGTTTTGCTGGTAATATTGTTCTGGGTGATATCATCGTT 180

Qy         79 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 98
               ||||||||||||   |||:::         :::   :::          |||:::   :::
Db        181 GCAGTGGACGGCAAACCTGTTAAGGGCAAATCTGACCTGCTGAGGGTTCTGGATGACTAT 240

Qy         99 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 118
                  ||||||   :::::::::|||                              |||   |||
Db        241 GGCGTCGGAGATCAGGTGACCTTGACA--------------------ATCCGGCGAGGC 279

Qy        119 AsnValThrLeu 122
               :::   ||||||
Db        280 TCAGAAACCCTT 291
```

RESULT 18
BI378928

LOCUS	BI378928 677 bp mRNA linear EST 26-AUG-2003
DEFINITION	BFLG1_000559 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMGp498) Branchiostoma floridae cDNA clone MPMGp498P1235 5', mRNA sequence.
ACCESSION	BI378928
VERSION	BI378928.1 GI:30914099
KEYWORDS	EST.
SOURCE	Branchiostoma floridae (Florida lancelet)
ORGANISM	Branchiostoma floridae Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.
REFERENCE	1 (bases 1 to 677)
AUTHORS	Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J., Herwig,R., Vingron,M. and Lehrach,H.
TITLE	New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes
JOURNAL	Genome Res. 13 (6A), 1056-1066 (2003)
MEDLINE	22683279
PUBMED	12799346
COMMENT	Contact: Panopoulou G laboratory 145, dept.Lehrach Max-Planck-Institut fuer Molekulare Genetik Ihnestr.63-73, D-14195 Berlin, Germany Tel: +49 30 8413 1235 Fax: +49 30 8413 1128 Email: panopoul@molgen.mpg.de The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: http://www.molgen.mpg.de/amphioxus/ Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (http://www.rzpd.de) PCR PRimers FORWARD: 5' CCCCAGGCTTTACACTTTATGCTTCCGGCTCG 3' (M13RSP) BACKWARD: 5' GCTATTACGCCAGCTGGCGAAAGGGGGATGTG 3' (M13FSP) Insert Length: 1200 Std Error: 0.00 Seq primer: 5'-CCGGTCCGGAATTCCCGGGT-3' pSport3/86 High quality sequence stop: 677.

FEATURES	Location/Qualifiers
source	1. .677 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="taxon:7739" /clone="MPMGp498P1235" /tissue_type="whole embryo" /dev_stage="5-6 hrs (gastrula stage)" /lab_host="E.coli, XL1 blue" /clone_lib="Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMGp498)" /note="Vector: pSport1; Site_1: SalI, KpnI, EcoRI (5'); Site_2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC (T)15-3' and a SalI 5'-TCGACCCACGCGTCCG-3'adapters (Gibco BRL)."

ORIGIN

Alignment Scores:

Pred. No.:	7.92	Length:	677
Score:	95.50	Matches:	21
Percent Similarity:	61.11%	Conservative:	12
Best Local Similarity:	38.89%	Mismatches:	18
Query Match:	14.62%	Indels:	3
DB:	12	Gaps:	1

US-09-684-215B-18 (1-128) x BI378928 (1-677)

```
Qy         54 GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThr 73
              |||      |||   |||:::::::      :::||||||   |||   ||||||
Db        143 GGAGTACTGGTGCACAGGATCATTGTGGGGTCACCTGCATACAGTGCGGGTATCCGTCCT 202

Qy         74 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 93
              ||||||||||||||:::::::|||      :::   |||||      :::   |||
Db        203 GGTGACGTCATCACCAGTATCAACGGTCGATAGGTGACGTCAGCACGAGACATCTACGAC 262

Qy         94 AlaLeuAsnGlyHisHisProGlyAspValIleSerValThr 107
              |||:::|||            |||      :::::::::|||
Db        263 GCGGTAAATAGC---------GGGCAGCAGCTGAACATCACC 295
```

RESULT 19
BI721127

LOCUS	BI721127 543 bp mRNA linear EST 19-SEP-2001
DEFINITION	1031054B10.y1 C. reinhardtii CC-1690, Stress II (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION	BI721127
VERSION	BI721127.1 GI:15696822
KEYWORDS	EST.
SOURCE	Chlamydomonas reinhardtii
ORGANISM	Chlamydomonas reinhardtii Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.
REFERENCE	1 (bases 1 to 543)
AUTHORS	Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
TITLE	Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1031
JOURNAL	Unpublished (2001)
COMMENT	Contact: Charles Hauser DCMB Box 91000 Duke University Durham, NC 27708-1000 Tel: 919 613 8159 Fax: 919 613 8177 Email: chauser@duke.edu.

FEATURES	Location/Qualifiers
source	1. .543 /organism="Chlamydomonas reinhardtii" /mol_type="mRNA" /strain="CC-1690 wild type mt+ 21gr" /db_xref="taxon:3055" /clone_lib="C. reinhardtii CC-1690, Stress II (normalized), Lambda Zap II" /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH4+ - containing) and shifted to TAP - NO3- (24hrs); H2 production conditions (0, 12hr, 24hr) see Melis et al.,(2000) Plant Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda

Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

ORIGIN

Alignment Scores:

Pred. No.:	6.41	Length:	543
Score:	95.00	Matches:	30
Percent Similarity:	40.71%	Conservative:	16
Best Local Similarity:	26.55%	Mismatches:	51
Query Match:	14.55%	Indels:	16
DB:	12	Gaps:	3

US-09-684-215B-18 (1-128) x BI721127 (1-543)

```
Qy         10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
              :::     |||   |||||||||:::||||||          :::      |||
Db        153 GCATCCTCCGGCGTGGGCTTCGCGCTGCCCATCGACAGTGTGCGCGGCCTGGTAGACCAG 212

Qy         30 IleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGly--------- 46
              |||           :::   |||       :::    |||
Db        213 ATCCTGACCTACGGCCGCGTGCTGCGGCCCGTGCTGGGAGTCAGCCTGGCGCCGCCGCAG 272

Qy         47 ValValAspAsnAsnGlyAsn---GlyAlaArgValGlnArgValValGlySerAlaPro 65
              |||:::         |||      |||      |||      |||       :::|||
Db        273 GTGCTCAAGCAGCTGGGCCAGCCGGGCGTGCTGGTGCTAGAGGTGCCCAAGGGCTCCCCC 332

Qy         66 AlaAlaSerLeuGlyIleSer----------------------------------Thr 73
              |||          ||||||
Db        333 GCGGAGAAGGCCGGCATCAAGCCCACCATGCGCGACCGCTTCAGCGGCTCCCTGGTGCTG 392

Qy         74 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 93
              ||||||:::||||||   :::||||||      :::   :::   :::   :::   :::
Db        393 GGCGACATCATCACAGGCATTGACGGCAAGGCCGTAAAGAACTATTCGGACCTGGTGGAG 452

Qy         94 AlaLeuAsnGlyHisHisProGlyAspValIleSerVal 106
              ||||||:::          ||||||   :::   |||
Db        453 GCGCTGGATGAGAAGCGCGTGGGCGACACAGTCAAGGTG 491
```

RESULT 20
BI956121

LOCUS	BI956121 673 bp mRNA linear EST 19-OCT-2001
DEFINITION	HVSMEm0025L21f Hordeum vulgare green seedling EST library HVcDNA0014 (Blumeria infected) Hordeum vulgare subsp. vulgare cDNA clone HVSMEm0025L21f, mRNA sequence.
ACCESSION	BI956121
VERSION	BI956121.1 GI:16303022
KEYWORDS	EST.
SOURCE	Hordeum vulgare subsp. vulgare
ORGANISM	Hordeum vulgare subsp. vulgare Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.
REFERENCE	1 (bases 1 to 673)
AUTHORS	Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Chin,A., Begum,D., Frisch,D., Atkins,M., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Oates,R. and Main,D.
TITLE	Development of a genetically and physically anchored EST resource for barley genomics: Blumeria infected Morex (compatible) seedling cDNA library
JOURNAL	Unpublished (2001)
COMMENT	Contact: Wing RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: rwing@clemson.edu Total hq bases = 479 Seq primer: AATTAACCCTCACTAAAGGG

High quality sequence start: 5
High quality sequence stop: 620.

FEATURES Location/Qualifiers

source 1. .673
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HVSMEm0025L21f"
/tissue_type="green seedling leaf"
/lab_host="TJC121"
/clone_lib="Hordeum vulgare green seedling EST library HVcDNA0014 (Blumeria infected)"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Morex (mla) plants were greenhouse grown in the R Wise lab at Iowa State University, Ames, IA; 7 day old green seedlings were infected with isolate 5874 of Blumeria graminis f. sp. hordei, and leaves were harvested 24, 48 and 72 hr post-inoculation and snap frozen (Wise). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids (Chin). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see http://www.genome.clemson.edu/projects/barley. To order this clone see http://www.genome.clemson.edu/orders Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

ORIGIN

Alignment Scores:

Pred. No.:	8.86	Length:	673
Score:	95.00	Matches:	39
Percent Similarity:	44.35%	Conservative:	16
Best Local Similarity:	31.45%	Mismatches:	49
Query Match:	14.55%	Indels:	20
DB:	12	Gaps:	6

US-09-684-215B-18 (1-128) x BI956121 (1-673)

```
Qy         10 SerGlnGlyGlyGln---GlyPheAlaIleProIle-----------------GlyGln 22
              :::|||||||||      |||||||||:::   :::                   |||
Db          6 AATCAAGGTGGAGTGTGAGGCTTCGCGATGGAGGTTGTTGAAACAGCGTATGCTGGTCAT 65

Qy         23 AlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThrAlaPhe 42
              |||     :::|||          :::   |||   |||          |||
Db         66 GCAAAGGTGCTTGCTTCTACTGTTGACCTCAGTACATGCCCTGATGGCATCATCTGTGTT 125

Qy         43 LeuGlyLeuGlyValValAspAsnAsnGlyAsnGly---------------------- 54
                 |||   |||||||||:::        ||||||
Db        126 GGGGGTGATGGAGTCGTAAATGAGGTTGTGAATGGTTTACTTGGGGGGGAGGATTTGAGA 185

Qy         55 AlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGly 74
                 :::|||       :::|||       ||||||   |||
Db        186 GAGGCGCTTCAACTTCCTATTGGGATAGTTCCTGCTGGTTCCGATAATTCATTAGTGTGG 245

Qy         75 AspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAla 94
                 |||:::         |||      |||:::::::::|||||||||:::|||
Db        246 AGTGTTCTTGGCATCAGGGAT------CCTGTGTCAGCAGCAACTGCTTTAGCCAAG--- 296
```

```
Qy         95 LeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGly 114
              |||       |||  ||||||  :::|||  |||      :::::|||
Db        297 ---GGTGGTTTCACACCGATTGATGTGTTTGCTGTAAAATGG---ATCCAAGCTGGAGTT 350

Qy        115 ThrArgThrGly 118
              |||  |||
Db        351 ACTCATTTCGGT 362

RESULT 21
BZ560689/c
LOCUS       BZ560689               1127 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  pacs2-164_2702.x1 pacs2-164 Pseudomonas aeruginosa genomic clone
            pacs2-164_2702, genomic survey sequence.
ACCESSION   BZ560689
VERSION     BZ560689.1  GI:27179214
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 1127)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .1127
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="2-164"
                     /db_xref="taxon:287"
                     /clone="pacs2-164_2702"
                     /clone_lib="pacs2-164"
                     /note="clinical isolate 2-164 Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:                21.8          Length:          1127
Score:                    94.50         Matches:         36
Percent Similarity:       41.32%        Conservative:    14
Best Local Similarity:    29.75%        Mismatches:      46
Query Match:              14.47%        Indels:          25
DB:                       28            Gaps:            6

US-09-684-215B-18 (1-128) x BZ560689 (1-1127)

Qy          13 GlyGlnGlyPheAlaIleProIleGly--------------GlnAlaMetAlaIleAla 27
               ||||||      |||   |||   |||                ||:::|||:::
Db         520 GGGCAGCACCACGCGCACCCAGCAGGAGGAGGTCTCGTCGTTCTTGCCGTCGCGGTCCCA 461

Qy          28 GlyGlnIleLysLeu---ProThrValHisIleGlyProThrAlaPheLeuGlyLeuGly 46
               |||::::::   |||   |||   |||   :::||||||             ||||||
Db         460 GGGGAACTGGATCTTGATCCGGCCGTACTGGTCGGTCCAGATCTCCTCGCCGGCCTTGCC 401

Qy          47 ValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAla 66
                  ::::::      |||       |||         :::|||       |||
Db         400 ---------CACCACCAGGGCGGTCTGGGTGTGCATGCGCGGCTTGGGCGTGACCCGCGC 350

Qy          67 AlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsn 86
               :::::::||      ||||||   :::   ||||||      |||   :::
Db         349 GGGGCGGTAGGCGGTGGCCTTGGGGATCGCCTCGAAGCGGTT------GCGGTAGTGCGC 296

Qy          87 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 106
```

```
                  |||   |||         |||   ||||||   |||:::      :::
Db         295 GTGACTGGCGTCATGGGTGAC------CGAGGTCAGCACCCAGTCGATATTCAGCGAATC 242

Qy         107 ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGluGlyPro 126
                                                   |||       |||||||||
Db         241 GTC-----------------------------------GTCATGCCCGCTGAGGGTGAA 218

Qy         127 Pro 127
               |||
Db         217 CCA 215

RESULT 22
CD080132
LOCUS       CD080132               982 bp    mRNA    linear   EST 14-SEP-2003
DEFINITION  MA3-9999U-M294-C03-U.B MA3-0001 Schistosoma mansoni cDNA clone
            MA3-9999U-M294-C03.B, mRNA sequence.
ACCESSION   CD080132
VERSION     CD080132.1  GI:34631128
KEYWORDS    EST.
SOURCE      Schistosoma mansoni
  ORGANISM  Schistosoma mansoni
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1  (bases 1 to 982)
  AUTHORS   Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M.,
            Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr.,
            Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F.,
            Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L.,
            Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A.,
            Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A.,
            Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T.,
            Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M.,
            Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
  TITLE     Transcriptome analysis of the acoelomate human parasite Schistosoma
            mansoni
  JOURNAL   Nat. Genet. 35 (2), 148-157 (2003)
  MEDLINE   22879926
COMMENT     Contact: Dr. Sergio Verjovski-Almeida
            Departamento de Bioquimica
            Instituto de Quimica - Universidade de Sao Paulo
            Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP,
            Brasil
            Tel: +55-11-3091-2173
            Fax: +55-11-3091-2186
            Email: verjo@iq.usp.br
            This sequence was derived from the FAPESP Schistosoma mansoni EST
            Genome Project. All sequences in the project were assembled and
            annotated. This entry and all the assembled sequences can be seen
            in the following URL http://bioinfo.iq.usp.br/schisto/
            Plate: MA3-9999U-M294  row: 3  column: C.
FEATURES             Location/Qualifiers
     source          1. .982
                     /organism="Schistosoma mansoni"
                     /mol_type="mRNA"
                     /db_xref="taxon:6183"
                     /clone="MA3-9999U-M294-C03.B"
                     /sex="mixed pool"
                     /dev_stage="adult"
                     /lab_host="Mus musculus"
                     /clone_lib="MA3-0001"
ORIGIN

Alignment Scores:
Pred. No.:                20            Length:          982
Score:                    94.00         Matches:         33
Percent Similarity:       35.94%        Conservative:    13
Best Local Similarity:    25.78%        Mismatches:      36
Query Match:              14.40%        Indels:          46
DB:                       14            Gaps:            5

US-09-684-215B-18 (1-128) x CD080132 (1-982)

Qy           6 AsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAla 25
```

```
Db        355 AACGCCATGGTGGCTGGTACAGGCGTGGGATTTGCTATACCTGTTGATCAAGTCCGAAAA 414

Qy         26 IleAlaGlyGlnIleLysLeu----------------------------ProThrVal 35

Db        415 TTCATA---CAGTTAGCACTAAAAGCATCCTCAAATACTCGGAGCTCTTCCCCAATATCG 471

Qy         36 HisIleGlyProThr------------------------------------------- 40

Db        472 TCAAAATCGCCTACTGACCCATATATTTTAGCAGACAATCCAGCCAATTCGGAAAGTGGA 531

Qy         41 ---------AlaPheLeuGlyLeu---------------------------------Gly 46

Db        532 ACTCAACGTCGATACTTAGGCCTAGTTATGCGGACACTTACTCCAGAACTGGCTTTCGAA 591

Qy         47 ValValAspAsnAsnGly-----------------AsnGlyAlaArgValGlnArgVal 60

Db        592 TTAGCATCTCGCGGTGGACAACATTTTGTAGAGTTGAATGGTGTTCTGATTCATGCTGTA 651

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80

Db        652 TTAAGAAACTCCCCTGCTCANAGAGCGGGATTAAGAGCTGGAGATGTAATTGTTGCCATT 711

Qy         81 AspGlyAlaProIleAsnSerAla 88

Db        712 GATGGTCTGCCTATAACGAATGCT 735

RESULT 23
CG322895/c
LOCUS       CG322895                771 bp    DNA     linear   GSS 26-AUG-2003
DEFINITION  OG0DK42TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0693G11,
            genomic survey sequence.
ACCESSION   CG322895
VERSION     CG322895.1  GI:34240161
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 771)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG0DK42TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0693G11"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

Alignment Scores:
Pred. No.:              15.7          Length:         771
Score:                  93.50         Matches:        40
Percent Similarity:     49.56%        Conservative:   16
Best Local Similarity:  35.40%        Mismatches:     45
Query Match:            14.32%        Indels:         13

DB:                     29            Gaps:           5

US-09-684-215B-18 (1-128) x CG322895 (1-771)

Qy         16 PheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrVal 35

Db        421 TTCGCCTTCTCCGCTTTCCGCTCGTCCGCGGCGGCGGGGACGACGGCATTGCCGACG--- 365

Qy         36 HisIleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsn--------------- 50

Db        364 AGCTTGGGATTGGCCCTGGGTTTAGCGCTCGGTGTCTCCGAGAGTACCTCCTTGACCGTT 305

Qy         51 AsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu--- 69

Db        304 TCCTCCTCCGGCGA-CGGCGGGTCGCGGTCCTCGGGTCCGCACCGGCGCGGGAAGTCAGA 246

Qy         70 ------GlyIleSerThrGlyAspValIleThrAlaValAspGlyAla-ProIleAsnSe 87

Db        245 GGCTCCGGCCACCGGGCGGGTCCTCGGGTTCTTGCT---GAAGCAGCAGCCCATTCTCCA 189

Qy         87 rAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGly---AspValIleSerVa 106

Db        188 GGCTCCAGCTGCCGCGGACGCACTAAGCAGCCATCGCGGTGGCCGTGATCTGGACTCTGT 129

Qy        106 lThrTrpGlnThrLysSerGlyGlyThrArgThrGly 118

Db        128 AGTGTGGCGTGTGGTGACCAGTGACAGTGAGAGGGGG 92

RESULT 24
CC634156/c
LOCUS       CC634156                787 bp    DNA     linear   GSS 19-JUN-2003
DEFINITION  OGOBL39TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0345G06,
            genomic survey sequence.
ACCESSION   CC634156
VERSION     CC634156.1  GI:32010854
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 787)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OGOBL39TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .787
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0345G06"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

Alignment Scores:
Pred. No.:              16.2          Length:         787
Score:                  93.50         Matches:        40
Percent Similarity:     49.56%        Conservative:   16
Best Local Similarity:  35.40%        Mismatches:     45
```

```
Query Match:          14.32%      Indels:        13
DB:                   29          Gaps:          5

US-09-684-215B-18 (1-128) x CC634156 (1-787)

Qy         16 PheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrVal 35
              ||||||           ::::::   |||   ||||||         |||||||||
Db        471 TTCGCCTTCTCCGCTTTCCGCTCGTCCGCGGCGGCGGGGACGACGGCATTGCCGACG--- 415

Qy         36 HisIleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsn--------------- 50
                 :::|||          |||   |||||||||   ::::::
Db        414 AGCTTGGGATTGGCCCTGGGTTTAGCGCTCGGTGTCTCCGAGAGTACCTCCTTGACCGTT 355

Qy         51 AsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu--- 69
              :::   :::|||   |||         |||:::|||||||||||||||   :::
Db        354 TCCTCCTCCGGCGA-CGGCGGGTCGCGGTCCTCGGGTCCGCACCGGCGCGGGAAGTCAGA 296

Qy         70 ------GlyIleSerThrGlyAspValIleThrAlaValAspGlyAla-ProIleAsnSe 87
                  |||          |||       :::   |||   :::   ||| ||||||
Db        295 GGCTCCGGCCACCGGGCGGGTCCTCGGGTTCTTGCT---GAAGCAGCAGCCCATTCTCCA 239

Qy         87 rAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGly---AspValIleSerVa 106
              |||   |||   |||||||||||||:::   |||       |||   |||:::   |||||
Db        238 GGCTCCAGCTGCCGCGGACGCACTAAGCAGCCATCGCGGTGGCCGTGATCTGGACTCTGT 179

Qy        106 lThrTrpGlnThrLysSerGlyGlyThrArgThrGly 118
              |   |||:::      :::       :::      |||
Db        178 AGTGTGGCGTGTGGTGACCAGTGACAGTGAGAGGGGG 142

RESULT 25
BZ567978/c
LOCUS       BZ567978                 885 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  pacs2-164_7306.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
            pacs2-164_7306, genomic survey sequence.
ACCESSION   BZ567978
VERSION     BZ567978.1  GI:27200212
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 885)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .885
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="2-164"
                     /db_xref="taxon:287"
                     /clone="pacs2-164_7306"
                     /clone_lib="pacs2-164"
                     /note="clinical isolate 2-164 Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:              19.3          Length:        885
Score:                  93.50         Matches:       37
Percent Similarity:     41.60%        Conservative:  15
Best Local Similarity:  29.60%        Mismatches:    48
Query Match:            14.32%        Indels:        25
DB:                     28            Gaps:          6

US-09-684-215B-18 (1-128) x BZ567978 (1-885)

Qy          9 LeuSerGlnGlyGlyGlnGlyPheAlaIleProIleGly---------------GlnAla 23
              |||:::      ||||||         |||   |||   |||                |||
Db        570 CTTGCCGCTCCAGGGCAGCACCACGCGCACCCAGCAGGAGGAGGTCTCGTCGTTCTTGCC 511

Qy         24 MetAlaIleAlaGlyGlnIleLysLeu---ProThrValHisIleGlyProThrAlaPhe 42
              :::|||:::   |||::::::   |||   |||   |||   :::||||||
Db        510 GTCGCGGTCCCAGGGGAACTGGATCTTGATCCGGCCGTACTGGTCGGTCCAGATCTCCTC 451

Qy         43 LeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGly 62
                ||||||           ::::::      |||       |||          :::|||
Db        450 GCCGGCCTTGCC---------CACCACCAGGTCGGTCTGGGTGTGCATGCGCGGCTTGGG 400

Qy         63 SerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGly 82
                  |||       ::::::|||          ||||||   :::   ||||||
Db        399 CGTGACCCGCGCGGGGCGGTAGGCGGTGGCCTTGGGGATCGCCTCGAAGCGGTT------ 346

Qy         83 AlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAsp 102
              |||   :::          |||   |||          |||   ||||||   |||
Db        345 GCGGTAGTGCGCGTGACTGGCGTCATGGGTGAC------CGAGGTCAGCACCCAGTCGAT 292

Qy        103 ValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 122
              :::      :::                                          |||
Db        291 ATTCAGCGAATCGTC----------------------------------GTCATGCCC 268

Qy        123 AlaGluGlyProPro 127
              |||||||||   |||
Db        267 GCTGAGGGTGAACCA 253

RESULT 26
CD383492
LOCUS       CD383492                 703 bp    mRNA    linear   EST 31-MAY-2003
DEFINITION  PTMM08663 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
            tricornutum cDNA 5', mRNA sequence.
ACCESSION   CD383492
VERSION     CD383492.1  GI:31259106
KEYWORDS    EST.
SOURCE      Phaeodactylum tricornutum
  ORGANISM  Phaeodactylum tricornutum
            Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
            Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
REFERENCE   1  (bases 1 to 703)
  AUTHORS   Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
  TITLE     Genome properties of the diatom Phaeodactylum tricornutum
  JOURNAL   Plant Physiol. 129 (3), 993-1002 (2002)
  MEDLINE   22111123
   PUBMED   12114555
COMMENT     Contact: Bowler C
            Laboratory of Molecular Plant Biology
            Stazione Zoologica 'Anton Dohrn'
            Villa Comunale, I-80121, Napoli, Italy
            Tel: 39 081 583 3268/3211
            Fax: 39 081 764 1355
            Email: chris@alpha.szn.it
            Diatom EST Database(http://avesthagen.sznbowler.com)
            Seq primer: T3 backward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .703
                     /organism="Phaeodactylum tricornutum"
                     /mol_type="mRNA"
                     /db_xref="taxon:2850"
                     /cell_line="CCMP632"
                     /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
                     /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
                     Xho I"
ORIGIN

Alignment Scores:
Pred. No.:              19.6          Length:        703
```

```
Score:                  92.00          Matches:       35
Percent Similarity:     40.00%         Conservative:  17
Best Local Similarity:  26.92%         Mismatches:    54
Query Match:            14.09%         Indels:        25
DB:                     14             Gaps:          4

US-09-684-215B-18 (1-128) x CD383492 (1-703)

Qy         10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla------ 27
              :::      |||   ||||||||||||||||||                  |||
Db        178 GCCTCGGCCGGTATTGGTTTTGCAATCCCAATTGACACCGTGAAATTTATTGTCGAGACG 237

Qy         28 ------------GlyGlnIleLysLeuProThrValHisIleGlyProThrAlaPhe--- 42
                         |||::::::     |||   :::   |||           :::
Db        238 CTCATTCGGGATGGCAGGGTGGTTCGACCAGTTTTGGGTATTTCCTACCTCGAATACAAA 297

Qy         43 ------LeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
                    ||||||:::          :::|||     |||
Db        298 CAAGCACGCGCTCTAGGGATA------------AATAGCGGTGTCTTGGTACTCGAAGCA 345

Qy         61 ValGlySerAlaProAlaAlaSerLeuGly---------------------------- 70
                  :::|||   |||      |||
Db        346 CCAGCCGGAAGCCCTCCTGCTCTCGCCGGTCTCAAAGGAACGCGACGAACCGAGTCTGGT 405

Qy         71 ---IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThr 89
                 :::      ||||||:::||||||   |||              |||      :::
Db        406 TTGGTAGAAATCGGTGACATTATAACGAAAGTGGGAGACAAGGTCATAACCGTGGAGTCC 465

Qy         90 AlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGln 109
              :::      ||||||      :::   |||||||||||||:::   ||||||   :::
Db        466 GATTTGTTTCAAGCTTTAGAAGAGTACAAGCCGGGTGATGTCGTGGACGTGACCGT-AAA 524

Qy        110 ThrLysSerGlyGlyThrArgThrGlyAsn 119
              :::      |||      |||:::   |||
Db        525 TCGGGTTTCGGCAGTGAACGATCAGCTAAC 554

RESULT 27
BG143271/c
LOCUS       BG143271                 600 bp    mRNA    linear   EST 13-MAR-2002
DEFINITION  ia95h02.y1 Melton Mouse E16 5 Pancreas Library 2 M16B2 Mus musculus
            cDNA clone IMAGE:5646194 5' similar to SW:PM1_HUMAN P17152 PUTATIVE
            RECEPTOR PROTEIN. ;, mRNA sequence.
ACCESSION   BG143271
VERSION     BG143271.1  GI:12646670
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Melton,D., Meadows,A., Clifton,S., Hillier,L., Marra,M., Pape,D.,
            Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B.,
            Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M.,
            McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and
            Bowers,Y.
  TITLE     WashU-Harvard Pancreas EST Project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Juliana Brown
            (brown@fas.harvard.edu)
            MGI:1852914 This sequence now available from the IMAGE consortium,
            for clone orders contact: info@image.llnl.gov
            Seq primer: -40RP from Gibco
            High quality sequence stop: 455.
FEATURES             Location/Qualifiers
     source          1. .600
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="ICR"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5646194"
                     /sex="Both"
                     /tissue_type="Total pancreas"
                     /dev_stage="Embryonic day 16.5"
                     /lab_host="TOP10"
                     /clone_lib="Melton Mouse E16 5 Pancreas Library 2 M16B2"
                     /note="Organ: Pancreas; Vector: pBluescript II SK; Site_1:
                     NotI; Site_2: SalI; Library constructed using SuperScript
                     Plasmid Library kit (Life  Technologies). cDNA made by
                     oligo-dT priming.Size-selected by  column fractionation;
                     average insert size 1.06kb. Primary  library,
                     unamplified."
ORIGIN

Alignment Scores:
Pred. No.:              17.4           Length:        600
Score:                  91.50          Matches:       37
Percent Similarity:     41.35%         Conservative:  18
Best Local Similarity:  27.82%         Mismatches:    45
Query Match:            14.01%         Indels:        33
DB:                     12             Gaps:          8

US-09-684-215B-18 (1-128) x BG143271 (1-600)

Qy          7 PheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIle 26
              :::   :::   ||||||   ||||||:::   :::                  ::::::
Db        573 TACACAGTACACCAGGGCAGCCAGGGCTATTGTGTT-----------GTGCAGTCTCTT 526

Qy         27 AlaGlyGlnIleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGly 46
              :::   |||:::   |||||||||   |||
Db        525 TCTGTGCAGGTCGTCCTTCCGACCCAGCAC---------------------------- 496

Qy         47 ValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAla 66
                           |||   |||   ||||||||||||      |||
Db        495 ------CACTGGTGTGGAGGAGGTGAGCGTGTGCAGAGGCAGGCGGGACAG--------- 451

Qy         67 AlaSerLeuGlyIleSerThrGlyAspValIleThrAla---------ValAspGlyAla 83
                    :::||||||              |||:::   |||          :::   |||
Db        450 ---TTTATAGGCATCATACTCCACCTGGTACTTGCAGCAAGGGTCAAACTGCCAGGAGAT 394

Qy         84 ProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspVal 103
              ||||||                  |||          :::|||          |||||||||
Db        393 CCCATAGAGGGTACAGCAGGCCAGGCTCAGCACACCAGCGGGCAG------GGAGATGTA 340

Qy        104 IleSerVal---------------ThrTrpGlnThrLysSerGlyGlyThr---ArgThr 117
              :::   |||                  |||      ::::::|||   |||   :::
Db        339 GTGGGAGTAATCCAGGGGCAGTGCCAATGGGGTGAAGAGGCAGGCAGTACCTGCCAACAC 280

Qy        118 GlyAsnVal------ThrLeuAlaGluGlyProProAla 128
              |||   :::       :::|||   |||   ||||||
Db        279 GGCCGTCTTGTGCAGGCAATTGCCTACGGTGATCCAGCG 241

RESULT 28
CD204316
LOCUS       CD204316                 701 bp    mRNA    linear   EST 20-MAY-2003
DEFINITION  HS1_7_H09.b1_A012 Heat-shocked seedlings Sorghum bicolor cDNA clone
            HS1_7_H09_A012 3', mRNA sequence.
ACCESSION   CD204316
VERSION     CD204316.1  GI:30933230
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 701)
  AUTHORS   Cordonnier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S.,
```

```
            Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Buchanan,C.D.,
            Eastman,A. and Pratt,L.H.
  TITLE     An EST database from Sorghum: heat-shocked seedlings
  JOURNAL   Unpublished (2003)
COMMENT     Other_ESTs: HS1_7_H09.g1_A012
            Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
            the Human Genome Center, University of Tokyo Institute of Medical
            Science; plant material and RNA prepared at Texas A & M University;
            sequencing done in the Laboratory for Genomics and Bioinformatics,
            University of Georgia. Sequence ends have been trimmed to exclude
            vector and regions below Phred quality 16. Three-prime sequences
            are presented as their reverse complement and have been trimmed to
            exclude polyA.
            Seq primer: Sug3 (CGACCTGCAGCTCGAGCACA)
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .701
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /cultivar="IS3620C"
                     /db_xref="taxon:4558"
                     /clone="HS1_7_H09_A012"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="Heat-shocked seedlings"
                     /note="Vector: pME18S-FL3; Site_1: XhoI; Site_2: XhoI; The
                     library was prepared from polyA+ RNA from 6-day-old
                     seedlings grown in hydroponic culture and heat-shocked at
                     40-42 C for 4 or 24 hr. After heat shock, roots and leaves
                     were harvested and tissues combined for RNA isolation.
                     Double-stranded cDNA was cloned unidirectionally into
                     different DraIII sites of the pME18S-FL3 vector (5-prime
                     DraIII site is CACTGTGTG, 3-prime DraIII site is
                     CACCATGTG)."
ORIGIN

Alignment Scores:
Pred. No.:                24.9          Length:         701
Score:                    91.00         Matches:        32
Percent Similarity:       38.71%        Conservative:   16
Best Local Similarity:    25.81%        Mismatches:     50
Query Match:              13.94%        Indels:         26
DB:                       14            Gaps:           4

US-09-684-215B-18 (1-128) x CD204316 (1-701)

Qy         16 PheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrVal 35
              ||||||:::||||||      :::   |||         |||
Db          3 TTTGCAGTACCGATTGACTCTGTCATTAAAATAGTAGAGAACTTTAAGAAAAATGGGAGA 62

Qy         36 HisIleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsn----------- 51
                 :::   |||      :::||||||||||   ::::::|||   |||
Db         63 GTTGTAAGGCCA------TGGCTTGGTTTAAAGATGCTTGACCTGAATCCCATGATCATT 116

Qy         52 -------------------------------------GlyAsnGlyAlaArgValGln 58
                                                     |||      |||
Db        117 GCACAGCTCAAAGAAAAATCAAGTACTTTTCCAGATGTAAGAAAAGGGGTGCTTGTTCCT 176

Qy         59 ArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThr 78
                 |||      ::::::||||||         |||         ||||||||||:::
Db        177 ATGGTTACACCAGCATCTCCAGCTGAACAAGCAGGATTTCGTCCTGGGGATGTGGTTGTT 236

Qy         79 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 98
                    |||   |||:::   |||        :::   |||   :::
Db        237 GAATTTGGTGGCAAACCAGTTGAGAGCATCAAAGAGATCATTGATATCATG--------- 287

Qy         99 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 118
                 ||||||   :::   |||   ::::::                    |||
Db        288 ------GGGGACAAGGTCGGAGTACCATTTAAAGTTCTTGTC------AAAAGAGCGAGC 335

Qy        119 AsnValThrLeu 122
                 |||||||||:::
Db        336 AATGTGACAGTG 347

RESULT 29
BU553475
LOCUS       BU553475                1098 bp    mRNA    linear   EST 16-SEP-2002
DEFINITION  AGENCOURT_10242287 NIH_MGC_109 Homo sapiens cDNA clone
            IMAGE:6577858 5', mRNA sequence.
ACCESSION   BU553475
VERSION     BU553475.1  GI:22903747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1098)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2778  row: k  column: 10
            High quality sequence start: 12
            High quality sequence stop: 552.
FEATURES             Location/Qualifiers
     source          1. .1098
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6577858"
                     /tissue_type="teratocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_109"
                     /note="Organ: ovary; Vector: pOTB7; Site_1: EcoRI; Site_2:
                     XhoI; cDNA made by oligo-dT priming.  Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G).  Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                49            Length:         1098
Score:                    91.00         Matches:        40
Percent Similarity:       39.85%        Conservative:   13
Best Local Similarity:    30.08%        Mismatches:     38
Query Match:              13.94%        Indels:         42
DB:                       13            Gaps:           8

US-09-684-215B-18 (1-128) x BU553475 (1-1098)

Qy         23 AlaMetAlaIleAlaGly----------GlnIleLysLeuProThrValHisIle 37
              |||   |||:::   |||              :::|||:::||||||
Db        340 GCTGCCGCGTTGGAAGGGCCATACCCAGTGGAGGAGATCCGTCTGCCAAGAGCTGGG--- 396

Qy         38 GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsn---------- 53
              ||||||         |||||||||   :::|||      :::   :::
Db        397 GGCCCT---------CTGGGGCTTACTATTGTCGGAGGCTCCGACCATTCCAGCCACCCG 447

Qy         54 -----------------GlyAlaArgValGlnArgValValGlySerAlaProAlaAla 67
```

```
                                  |||      :::   :::|||:::            ||||||
Db        448 TTTGGTGTCCAGGAGCCTGGCGTGTTCATCTCCAAGGTGCTCCCGCGGGGCCTGGCCGCT 507

Qy         68 SerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSer 87
              |||:::      ||||||   |||   ||||||:::|||      :::
Db        508 CGCAGCGGCCTGCGGGTTGGGGACCGCATCCTGGCCGTGAACGGGCAAGACGTGCGGGAT 567

Qy         88 AlaThr-----------------AlaMetAlaAspAlaLeu----------------- 95
              ||||||                 |||:::      ||||||
Db        568 GCCACCCACCAAGAAACAGTCAGTGCCCTGCTCCCGGCCCTGCCCTGGGAGCTTGTCGCT 627

Qy         96 -----------------AsnGlyHisHisProGlyAspValIleSerValThrTrpGln 109
                               ||||||   ||||||               ||||||
Db        628 GCTGGGTGCGCGAGGGGACCCGGGCACCCCCCCGGGCCC---TACGGGCAAACCTGGTGC 684

Qy        110 ThrLysSer------------GlyGlyThrArgThrGly 118
              |||                 ||||||||||||   |||
Db        685 ACTCCCCCAAAAGGGACCCTTGGCGGAACAAGGCCTGGG 723
```

RESULT 30
CF512107/c
LOCUS CF512107 631 bp mRNA linear EST 09-SEP-2003
DEFINITION CAbud0004_IVR_E10 Vitis vinifera cv. cabernet sauvignon (Clone 8) Bud - CABUD Vitis vinifera cDNA clone CAbud0004_IVR_E10 3', mRNA sequence.
ACCESSION CF512107
VERSION CF512107.1 GI:34543875
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis.
REFERENCE 1 (bases 1 to 631)
AUTHORS Goes da Silva,F., Iandolino,A., Lim,H., Baek,J., Jones,K. and Cook,D.
TITLE Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon' berries at various developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq primer: GCCAAACGAATGGTCTAG.
FEATURES Location/Qualifiers
source 1. .631
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon (Clone 8)"
/db_xref="taxon:29760"
/clone="CAbud0004_IVR_E10"
/sex="Hermaphrodite"
/dev_stage="Pre-bloom (10-11 days before bloom)"
/lab_host="DH5alpha"
/clone_lib="Vitis vinifera cv. cabernet sauvignon (Clone 8) Bud - CABUD"
/note="Organ: Bud; Vector: pDNR; Site_1: SfiI; Site_2: SfiI; CABUD is a cDNA library of Vitis vinifera cv. 'Cabernet Sauvignon' Clone 8 dissected buds. Samples were collected May 13, 2002 from pre-bloom plants (10-11 days before bloom), pre-veraison. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5'and 3' adaptors were used in cloning as follows:
5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."
ORIGIN

Alignment Scores:

Pred. No.:	23.9	Length:	631
Score:	90.50	Matches:	31
Percent Similarity:	44.44%	Conservative:	21
Best Local Similarity:	26.50%	Mismatches:	40
Query Match:	13.86%	Indels:	25
DB:	14	Gaps:	5

US-09-684-215B-18 (1-128) x CF512107 (1-631)

```
Qy         10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
              :::      |||  ||||||:::||||||:::                |||     |||
Db        574 GCATCTTCTGGTGTCGGATTTTCAATTCCAGTTGACACTGTAAGCGGCATTGTCGACCAA 515

Qy         30 Ile-----------LysLeuProThrValHisIleGly-----------ProThrAla 41
              :::            |||:::|||         :::|||            |||
Db        514 TTGGTGAGGTTTGGAAAAGTCCCAAGACCTATTTTAGGAATTAAGTTTGCGCCTGATCAG 455

Qy         42 PheLeuGly---LeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              :::|||   |||||||||                :::||||||   |||
Db        454 TCTGTGGGGCAGTTGGGTGTA--------------AGTGGGGCCCTTGTTTTAGATGCT 410

Qy         61 ValGlySerAlaProAlaAlaSerLeuGly---------------------------- 70
              :::  ||||||   |||
Db        409 CCTGCAAACGGTCCTGCTGGCAAAGCGGGCCTACTACCAACCAAACGTGATGCCTATGGC 350

Qy         71 ---IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThr 89
              :::      ||||||:::||||||:::|||:::|||      :::::::::   :::
Db        349 AGACTTATTTTAGGTGACATCATAACATCGGTGAATGGAAAAAAGGTTTCCAATGGAAGC 290

Qy         90 AlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 106
              :::      |||:::      ||||||   ::::::|||
Db        289 GACTTGTACAGAATTCTTGACCAGTGTAAAGTGGGTGACACGGTGACTGTG 239
```

Search completed: April 30, 2004, 04:40:59
Job time : 2513.02 secs

```
RESULT 30
AX832594
LOCUS       AX832594               1872 bp    DNA     linear   PAT 12-DEC-2003
DEFINITION  Sequence 17 from Patent EP1347055.
ACCESSION   AX832594
VERSION     AX832594.1  GI:39840644
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Reed,S.G., Skeiky,Y.A., Campos-Neto,A., Houghton,R.L.,
            Vedvick,T.S., Twardzik,D.R. and Dillon,D.C.
  TITLE     Compounds for immunotherapy and diagnosis of tuberculosis
  JOURNAL   Patent: EP 1347055-A 17 24-SEP-2003;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:              2.33e-41        Length:         1872
Score:                  646.00          Matches:        127
Percent Similarity:     99.22%          Conservative:   0
Best Local Similarity:  99.22%          Mismatches:     1
Query Match:            99.23%          Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-23 (1-128) x AX832594 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAG 1141


Search completed: April 30, 2004, 02:25:07
Job time : 3732.05 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 22:43:07 ; Search time 469.129 Seconds
(without alignments)
1234.209 Million cell updates/sec

Title: US-09-684-215B-18
Perfect score: 653
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....QTKSGGTRTGNVTLAEGPPA 128

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061306_13266/app_query.fasta_1.1180
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=30 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09684215_@CGN_1_1_370_@runat_29042004_061306_13266
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	653	100.0	675	9	US-09-759-143-822	Sequence 822, App
2	653	100.0	675	9	US-09-780-669-822	Sequence 822, App
3	653	100.0	675	9	US-09-822-827-822	Sequence 822, App
4	653	100.0	675	9	US-09-895-793-822	Sequence 822, App
5	653	100.0	675	9	US-09-895-814-822	Sequence 822, App
6	653	100.0	675	14	US-10-012-896-822	Sequence 822, App
7	653	100.0	675	15	US-10-144-678A-822	Sequence 822, App
8	653	100.0	675	15	US-10-294-025-822	Sequence 822, App
9	653	100.0	822	9	US-09-736-457-1862	Sequence 1862, Ap
10	653	100.0	822	9	US-09-902-941-1862	Sequence 1862, Ap
11	653	100.0	822	9	US-09-849-626-1862	Sequence 1862, Ap
12	653	100.0	822	13	US-10-283-017-1862	Sequence 1862, Ap
13	653	100.0	822	15	US-10-017-754-1862	Sequence 1862, Ap
14	653	100.0	822	15	US-10-113-872-1862	Sequence 1862, Ap
15	653	100.0	894	13	US-09-878-722-236	Sequence 236, App
16	653	100.0	894	13	US-09-904-456-236	Sequence 236, App
17	653	100.0	900	9	US-09-735-705-353	Sequence 353, App
18	653	100.0	900	9	US-09-850-716A-353	Sequence 353, App
19	653	100.0	900	9	US-09-897-778-353	Sequence 353, App
20	653	100.0	900	13	US-10-007-700-353	Sequence 353, App
21	653	100.0	900	15	US-10-117-982-353	Sequence 353, App
22	653	100.0	900	16	US-10-313-986-353	Sequence 353, App
23	653	100.0	915	9	US-09-759-143-834	Sequence 834, App
24	653	100.0	915	9	US-09-780-669-834	Sequence 834, App
25	653	100.0	915	9	US-09-822-827-834	Sequence 834, App
26	653	100.0	915	9	US-09-895-793-834	Sequence 834, App
27	653	100.0	915	9	US-09-895-814-834	Sequence 834, App
28	653	100.0	915	14	US-10-012-896-834	Sequence 834, App
29	653	100.0	915	15	US-10-144-678A-834	Sequence 834, App
30	653	100.0	915	15	US-10-294-025-834	Sequence 834, App
31	653	100.0	945	9	US-09-736-457-1861	Sequence 1861, Ap
32	653	100.0	945	9	US-09-902-941-1861	Sequence 1861, Ap
33	653	100.0	945	9	US-09-849-626-1861	Sequence 1861, Ap
34	653	100.0	945	13	US-10-283-017-1861	Sequence 1861, Ap
35	653	100.0	945	15	US-10-017-754-1861	Sequence 1861, Ap
36	653	100.0	945	15	US-10-113-872-1861	Sequence 1861, Ap
37	653	100.0	1012	9	US-09-735-705-351	Sequence 351, App
38	653	100.0	1012	9	US-09-850-716A-351	Sequence 351, App
39	653	100.0	1012	9	US-09-897-778-351	Sequence 351, App
40	653	100.0	1012	13	US-10-007-700-351	Sequence 351, App
41	653	100.0	1012	15	US-10-117-982-351	Sequence 351, App
42	653	100.0	1012	16	US-10-313-986-351	Sequence 351, App
43	653	100.0	1035	9	US-09-922-217-1084	Sequence 1084, Ap
44	653	100.0	1035	9	US-09-833-263-1084	Sequence 1084, Ap
45	653	100.0	1035	10	US-09-938-864-388	Sequence 388, App

ALIGNMENTS

RESULT 1
US-09-759-143-822
; Sequence 822, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Panger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 29, 2004, 23:03:22 ; Search time 291 Seconds
(without alignments)
5781.053 Million cell updates/sec

Title: US-09-684-215B-3
Perfect score: 396
Sequence: 1 acggccgcgtccgataactt.....tggccgagggacccccggcc 396

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	396	100.0	396	4	AAL40769	Aal40769 Nucleotid
2	396	100.0	672	4	AAL40772	Aal40772 Nucleotid
3	396	100.0	702	2	AAZ20206	Aaz20206 Mycobacte
4	396	100.0	702	4	AAL40770	Aal40770 Nucleotid
5	396	100.0	702	6	ABK14140	Abk14140 DNA encod
6	396	100.0	1002	6	AAD47078	Aad47078 Mycobacte
7	396	100.0	1002	6	AAD47077	Aad47077 Mycobacte
8	396	100.0	1002	6	AAD28336	Aad28336 Mycobacte
9	396	100.0	1002	6	AAD28337	Aad28337 Mycobacte
10	396	100.0	1068	2	AAX34251	Aax34251 Mycobacte
11	396	100.0	1143	2	AAX34252	Aax34252 Mycobacte
12	396	100.0	1742	4	AAL40771	Aal40771 Nucleotid
13	396	100.0	2181	8	ADA26353	Ada26353 Mycobacte
14	396	100.0	2190	6	AAD47084	Aad47084 Mycobacte
15	396	100.0	2190	6	AAD28343	Aad28343 Mycobacte
16	396	100.0	2191	4	AAL40773	Aal40773 Nucleotid
17	396	100.0	2286	6	ABK14128	Abk14128 DNA encod
18	396	100.0	2287	2	AAZ20194	Aaz20194 Mycobacte
19	396	100.0	2287	6	AAD47083	Aad47083 Mycobacte
20	396	100.0	2287	6	AAD28342	Aad28342 Mycobacte
21	396	100.0	2451	8	ADA26360	Ada26360 Mycobacte
22	396	100.0	2487	8	ADA26359	Ada26359 Mycobacte
23	396	100.0	2637	8	ADA26358	Ada26358 Mycobacte
24	396	100.0	2808	6	AAD47110	Aad47110 Mycobacte
25	396	100.0	2808	8	ADA26357	Ada26357 Mycobacte
26	396	100.0	3030	8	ADA26355	Ada26355 Mycobacte
27	396	100.0	3060	8	ADA26363	Ada26363 M. bovis
28	396	100.0	3104	8	ADA26362	Ada26362 Mycobacte
29	396	100.0	3474	8	ADA26361	Ada26361 Mycobacte
30	396	100.0	110000	4	AAI99682_01	Continuation (2 of
31	396	100.0	110000	4	AAI99683_01	Continuation (2 of
32	394.4	99.6	447	2	AAT91403	Aat91403 Mycobacte
33	394.4	99.6	447	2	AAT91466	Aat91466 Mycobacte
34	394.4	99.6	447	2	AAV44342	Aav44342 Mycobacte
35	394.4	99.6	447	2	AAV64450	Aav64450 M. tuberc
36	394.4	99.6	447	2	AAZ19040	Aaz19040 M. tuberc
37	394.4	99.6	447	2	AAZ19252	Aaz19252 M. tuberc
38	394.4	99.6	447	5	AAS03780	Aas03780 M. tuberc
39	394.4	99.6	447	6	AAD47080	Aad47080 Mycobacte
40	394.4	99.6	447	6	AAD28339	Aad28339 Mycobacte
41	394.4	99.6	1629	5	AAS03793	Aas03793 M. tuberc
42	392.8	99.2	1871	4	AAL40768	Aal40768 Nucleotid
43	391.2	98.8	1872	2	AAT91414	Aat91414 Mycobacte
44	391.2	98.8	1872	2	AAT91477	Aat91477 Mycobacte
45	391.2	98.8	1872	2	AAV44355	Aav44355 Mycobacte

ALIGNMENTS

RESULT 1
AAL40769
ID AAL40769 standard; DNA; 396 BP.
XX
AC AAL40769;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ra12 protein.
XX
KW Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1. .396
FT /*tag= a
FT /product= "Ra12 protein"
FT /note= "No start or stop codon"
XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027652.
XX
PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J;
XX
DR WPI; 2001-266299/27.
DR P-PSDB; AAO22138.
XX
PT Recombinant nucleic acid molecule for producing high yield expression of
PT desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS Claim 1; Fig 2; 39pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of

```
CC   Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC   The recombinant fusion nucleic acids and polypeptides are useful for
CC   providing stable and high yield expression of fusion polypeptides of both
CC   eukaryotic and prokaryotic origin and to encode a protein product for use
CC   as an antigen for detecting serum antibodies. The presence of serum
CC   antibodies to M. tuberculosis antigens in an individual indicates that
CC   the individual is infected with it. The fusion polypeptides are useful as
CC   sources of proteins for monitoring binding of serum antibodies to fusion
CC   proteins and as an immunogen to induce and/or enhance immune responses.
CC   The coding sequences can be ligated with a coding sequence of another
CC   molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC   can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC   represents the DNA encoding the Ra12 protein
XX
SQ   Sequence 396 BP; 68 A; 132 C; 134 G; 62 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 4;  Length 396;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396


RESULT 2
AAL40772
ID   AAL40772 standard; DNA; 672 BP.
XX
AC   AAL40772;
XX
DT   06-AUG-2003  (revised)
DT   03-OCT-2002  (first entry)
XX
DE   Nucleotide sequence encoding Ra12-mammaglobin fusion protein.
XX
KW   Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW   vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW   immunogen; cytokine; gene; ds.
XX
OS   Mammalia.
OS   Unidentified.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .666
FT                   /*tag= a
FT                   /product= "Ra12-mammaglobin fusion protein"
XX
```

```
PN   WO200125401-A2.
XX
PD   12-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US027652.
XX
PR   07-OCT-1999;   99US-0158585P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J;
XX
DR   WPI; 2001-266299/27.
DR   P-PSDB; AAO22141.
XX
PT   Recombinant nucleic acid molecule for producing high yield expression of
PT   desired fusion polypeptides, encodes fusion polypeptide comprising
PT   Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS   Disclosure; Fig 5; 39pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC   kDa C-terminal fragment of serine protease antigen MTB32A of
CC   Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC   The recombinant fusion nucleic acids and polypeptides are useful for
CC   providing stable and high yield expression of fusion polypeptides of both
CC   eukaryotic and prokaryotic origin and to encode a protein product for use
CC   as an antigen for detecting serum antibodies. The presence of serum
CC   antibodies to M. tuberculosis antigens in an individual indicates that
CC   the individual is infected with it. The fusion polypeptides are useful as
CC   sources of proteins for monitoring binding of serum antibodies to fusion
CC   proteins and as an immunogen to induce and/or enhance immune responses.
CC   The coding sequences can be ligated with a coding sequence of another
CC   molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC   can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC   represents the DNA encoding the Ra12-mammaglobin fusion protein. (Updated
CC   on 06-AUG-2003 to correct OS field.)
XX
SQ   Sequence 672 BP; 160 A; 185 C; 187 G; 140 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 4;  Length 672;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

```
RESULT 3
AAZ20206
ID   AAZ20206 standard; DNA; 702 BP.
XX
AC   AAZ20206;
XX
DT   17-JAN-2000  (first entry)
XX
DE   Mycobacterium tuberculosis antigen fusion protein Mtb24 DNA.
XX
KW   Tuberculosis; antigen; fusion protein; Mtb24; Ra12; DPPD; diagnosis;
KW   therapy; vaccine; immunogen; ss.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO9951748-A2.
XX
PD   14-OCT-1999.
XX
PF   07-APR-1999;   99WO-US007717.
XX
PR   07-APR-1998;   98US-00056556.
PR   30-DEC-1998;   98US-00223040.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Alderson M,  Campos-Neto A;
XX
DR   WPI; 1999-601610/51.
DR   P-PSDB; AAY32071.
XX
PT   New fusion proteins useful for diagnosis, prevention and treatment of
PT   tuberculosis.
XX
PS   Example; Fig 13A-B; 83pp; English.
XX
CC   This DNA sequence includes a coding region for a recombinant
CC   Mycobacterium tuberculosis bi-antigen fusion protein (see AAY32071),
CC   termed Mtb24, composed of the antigens Ra12 and DPPD. The DNA is useful
CC   for the recombinant production of the fusion protein. Coding sequences
CC   for the antigens were modified by PCR in order to facilitate their fusion
CC   and subsequent expression of the fusion protein, and then ligated. The
CC   invention provides fusion proteins (see AAY32059-71) containing at least
CC   2 M. tuberculosis antigens. The new fusion proteins and polynucleotides
CC   encoding them are useful as vaccines for preventing tuberculosis
CC   (claimed), for diagnosis (via in vitro assays or intradermal skin tests
CC   for detection of anti-M. tuberculosis antibodies), monitoring of disease
CC   progression, and treatment of tuberculosis. They are more effective
CC   immunogens than mixtures of the individual protein components
XX
SQ   Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 2;  Length 702;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 4
AAL40770
ID   AAL40770 standard; DNA; 702 BP.
XX
AC   AAL40770;
XX
DT   03-OCT-2002  (first entry)
XX
DE   Nucleotide sequence encoding Ra12-DPPD fusion protein.
XX
KW   Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW   vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW   immunogen; cytokine; gene; ds.
XX
OS   Unidentified.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .696
FT                   /*tag=  a
FT                   /product= "Ra12-DPPD fusion protein"
XX
PN   WO200125401-A2.
XX
PD   12-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US027652.
XX
PR   07-OCT-1999;   99US-0158585P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J;
XX
DR   WPI; 2001-266299/27.
DR   P-PSDB; AAO22139.
XX
PT   Recombinant nucleic acid molecule for producing high yield expression of
PT   desired fusion polypeptides, encodes fusion polypeptide comprising
PT   Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS   Example 1; Fig 3; 39pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC   kDa C-terminal fragment of serine protease antigen MTB32A of
CC   Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC   The recombinant fusion nucleic acids and polypeptides are useful for
CC   providing stable and high yield expression of fusion polypeptides of both
CC   eukaryotic and prokaryotic origin and to encode a protein product for use
CC   as an antigen for detecting serum antibodies. The presence of serum
CC   antibodies to M. tuberculosis antigens in an individual indicates that
CC   the individual is infected with it. The fusion polypeptides are useful as
CC   sources of proteins for monitoring binding of serum antibodies to fusion
CC   proteins and as an immunogen to induce and/or enhance immune responses.
CC   The coding sequences can be ligated with a coding sequence of another
CC   molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC   can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC   represents the DNA encoding the Ra12-DPPD fusion protein
XX
```

```
SQ   Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 4;  Length 702;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 5
ABK14140
ID   ABK14140 standard; DNA; 702 BP.
XX
AC   ABK14140;
XX
DT   29-AUG-2003  (revised)
DT   08-MAY-2002  (first entry)
XX
DE   DNA encoding antigenic fusion protein Ra12-DPPD (Mtb24).
XX
KW   Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;
KW   tuberculostatic; immunogen; vaccine; Ra12-DPPD; Mtb24.
XX
OS   Mycobacterium tuberculosis.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .702
FT                   /*tag=  a
FT                   /product= "Mtb24 #1"
FT                   /partial
FT                   /transl_except= (pos:694. .696, aa:Xaa)
FT                   /note= "No start or stop codon. Xaa= In frame stop codon"
FT   CDS             2. .702
FT                   /*tag=  b
FT                   /product= "Mtb24 #2"
FT                   /partial
FT                   /note= "No start or stop codon"
FT                   /transl_except= (pos:263. .265, aa:Xaa) /transl_except=
FT                   (pos:353. .355, aa:Xaa) /transl_except= (pos:395. .397,
FT                   aa:Xaa) /transl_except= (pos:470. .472, aa:Xaa)
FT                   /transl_except= (pos:701. .702, aa:Ser)
FT                   /note= "This codon has an apparent 1 nucleotide deletion
FT                   which alters the reading frame. Xaa= In frame stop codon"
FT   CDS             3. .701
FT                   /*tag=  c
FT                   /product= "Mtb24 #3"
FT                   /partial
FT                   /note= "No start or stop codon"
FT                   /transl_except= (pos:1. .2, aa:Pro) /transl_except=
FT                   (pos:39. .41, aa:Xaa) /transl_except= (pos:321. .323,
FT                   aa:Xaa) /transl_except= (pos:339. .341, aa:Xaa)
FT                   /transl_except= (pos:450. .452, aa:Xaa) /transl_except=
FT                   (pos:621. .623, aa:Xaa)
FT                   /note= "No start or stop codon. Xaa= In frame stop codon"
XX
PN   US2002009459-A1.
XX
PD   24-JAN-2002.
XX
PF   07-APR-1999;   99US-00287849.
XX
PR   13-MAR-1997;   97US-00818112.
PR   01-OCT-1997;   97US-00942578.
PR   18-FEB-1998;   98US-00025197.
PR   07-APR-1998;   98US-00056556.
PR   30-DEC-1998;   98US-00223040.
XX
PA   (REED/) REED S G.
PA   (SKEI/) SKEIKY Y A.
PA   (DILL/) DILLON D C.
PA   (ALDE/) ALDERSON M.
PA   (CAMP/) CAMPOS-NETO A.
XX
PI   Reed SG,  Skeiky YA,  Dillon DC,  Alderson M,  Campos-Neto A;
XX
DR   WPI; 2002-171134/22.
DR   P-PSDB; AAU74600, AAU76541, AAU76542.
XX
PT   New fusion proteins of Mycobacterium tuberculosis antigens, useful for
PT   diagnosing, treating or preventing M. tuberculosis infection,
PT   particularly as vaccine for treating or preventing tuberculosis.
XX
PS   Example; Fig 13; 62pp; English.
XX
CC   The invention relates to a purified polypeptide which induces an immune
CC   response of Mycobacterium tuberculosis. Polypeptides of the invention are
CC   useful for diagnosing, treating or preventing M. tuberculosis infection,
CC   particularly tuberculosis infection. In particular, the polypeptides are
CC   useful as a vaccine formulation with an adjuvant to afford long-term
CC   protection in animals against the development of tuberculosis. The
CC   protein coding sequence may be used to encode a protein product for use
CC   as an immunogen to induce and/or enhance an immune response to M.
CC   tuberculosis. This sequence represents DNA encoding an M. tuberculosis
CC   fusion protein of the invention. This polynucleotide encodes 3 different
CC   proteins, each in a different reading frame. (Updated on 29-AUG-2003 to
CC   standardise OS field)
XX
SQ   Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 6;  Length 702;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264
```

```
Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 6
AAD47078
ID AAD47078 standard; DNA; 1002 BP.
XX
AC AAD47078;
XX
DT 27-JAN-2003 (first entry)
XX
DE Mycobacterium tuberculosis Ra35FLMutSA mutant antigenic protein DNA.
XX
KW Vaccine; immunity; diagnostic agent; gene therapy; Ra35FLMutSA; mutant;
KW gene; antigen; ds.
XX
OS Mycobacterium tuberculosis.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 4. .996
FT /*tag= a
FT /product= "Ra35FLMutSA mutant antigenic protein"
XX
PN WO200272792-A2.
XX
PD 19-SEP-2002.
XX
PF 13-MAR-2002; 2002WO-US008223.
XX
PR 13-MAR-2001; 2001US-0275837P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Brannon M, Guderian J;
XX
DR WPI; 2002-759844/82.
DR P-PSDB; AAE29703.
XX
PT New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT tuberculosis.
XX
PS Disclosure; Page 80-81; 155pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC polynucleotide sequence encoding an antigen or an antigenic fragment from
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC polypeptide or its fragment. The Leishmania polynucleotide is selected
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides, as
CC in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is M.
CC tuberculosis Ra35FLMutSA mutant antigenic protein encoding DNA
XX
SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 6; Length 1002;
Best Local Similarity 100.0%; Pred. No. 5.6e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993
```

RESULT 7
AAD47077
ID AAD47077 standard; DNA; 1002 BP.
XX
AC AAD47077;
XX
DT 27-JAN-2003 (first entry)
XX
DE Mycobacterium tuberculosis mature Ra35 antigen encoding DNA.
XX
KW Vaccine; immunity; diagnostic agent; gene therapy; Ra35 antigen; gene;
KW ds.
XX
OS Mycobacterium tuberculosis.
XX
FH Key Location/Qualifiers
FT CDS 4. .996
FT /*tag= a
FT /product= "Ra35 mature antigenic protein"
FT /transl_except= (pos:547. .549, aa:Asp)
FT /transl_except= (pos:550. .552, aa:Ser)
XX
PN WO200272792-A2.
XX
PD 19-SEP-2002.
XX
PF 13-MAR-2002; 2002WO-US008223.
XX
PR 13-MAR-2001; 2001US-0275837P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Brannon M, Guderian J;
XX
DR WPI; 2002-759844/82.
DR P-PSDB; AAE29702.
XX
PT New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT tuberculosis.

```
XX
PS   Disclosure; Page 79; 155pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC   polynucleotide sequence encoding an antigen or an antigenic fragment from
CC   Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC   polypeptide or its fragment. The Leishmania polynucleotide is selected
CC   from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC   are used in methods for eliciting immune response in mammals. They are
CC   useful as vaccines to elicit protective immunity against pathogenic
CC   microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC   polypeptides are used for enhancing the expression of polynucleotides, as
CC   in vivo diagnostic agents and for raising antibodies in a non-human
CC   animal. The invention is used in gene therapy. The present sequence is M.
CC   tuberculosis mature Ra35 (N-terminus of MTB32A; RA35FL) antigen encoding
CC   DNA
XX
SQ   Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 6;  Length 1002;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993


RESULT 8
AAD28336
ID   AAD28336 standard; cDNA; 1002 BP.
XX
AC   AAD28336;
XX
DT   22-APR-2002  (first entry)
XX
DE   Mycobacterium sp. MTB32A (Ra35FL) mature protein encoding cDNA.
XX
KW   Fusion protein; antigen; serological sensitivity; immune response;
KW   tuberculosis; infection; vaccine; MTB32A; Ra32FL; ss.
XX
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .996
FT                   /*tag=  a
FT                   /product= "Ra35 mature protein"
FT                   /transl_except= (pos:547. .549, aa:Asp)
FT                   /transl_except= (pos:550. .552, aa:Ser)
XX
PN   WO200198460-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019959.
XX
PR   20-JUN-2000; 2000US-00597796.
PR   01-FEB-2001; 2001US-0265737P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Reed S,  Alderson M;
XX
DR   WPI; 2002-147798/19.
DR   P-PSDB; AAE17566.
XX
PT   Composition comprising MTB39 antigen and MTB32A antigen from
PT   Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS   Disclosure; Page 95; 136pp; English.
XX
CC   The present invention relates to fusion proteins containing at least two
CC   Mycobacterium species antigens, nucleotides encoding them and
CC   compositions comprising such fusion proteins. The present invention
CC   particularly relates to nucleic acids encoding fusion proteins that
CC   include two or more individual M. tuberculosis antigens which increase
CC   the serological sensitivity of sera from individuals infected with
CC   tuberculosis and methods for their use in diagnosis, prevention and
CC   treatment of tuberculosis infection. Sequences of the invention are
CC   useful for eliciting an immune response in a mammal, e.g., human,
CC   immunised with BCG. They are useful in the diagnosis, treatment and
CC   prevention of Mycobacterium infection. The fusion proteins and the
CC   polynucleotides are useful as diagnostic tools in patients infected with
CC   Mycobacterium, in vitro and in vivo assays for detecting humoral
CC   antibodies or cell-mediated immunity against M. tuberculosis, for the
CC   diagnosis of an infection or monitoring of disease progression, as
CC   immunogens to generate or elicit a protective immune response in a
CC   patient and for raising anti-M. tuberculosis antibodies in a non-human
CC   animal. Sequences of the invention are also used as vaccines. MTB32A
CC   fusion proteins of the invention are useful as in vivo diagnostic agents
CC   for intradermal skin test. The present sequence is a cDNA encoding
CC   Mycobacterium species MTB32A (Ra32FL) mature protein
XX
SQ   Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 6;  Length 1002;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993
```

RESULT 9
AAD28337
ID AAD28337 standard; cDNA; 1002 BP.
XX
AC AAD28337;
XX
DT 22-APR-2002 (first entry)
XX
DE Mycobacterium species Ra35FLMutSA mutant cDNA.
XX
KW Fusion protein; antigen; serological sensitivity; immune response;
KW tuberculosis; infection; vaccine; Ra32FLMutSA; mutant; mutein; ss.
XX
OS Mycobacterium sp.
XX
FH Key Location/Qualifiers
FT CDS 4. .996
FT /*tag= a
FT /product= "Ra35FLMutSA protein"
XX
PN WO200198460-A2.
XX
PD 27-DEC-2001.
XX
PF 20-JUN-2001; 2001WO-US019959.
XX
PR 20-JUN-2000; 2000US-00597796.
PR 01-FEB-2001; 2001US-0265737P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Reed S, Alderson M;
XX
DR WPI; 2002-147798/19.
DR P-PSDB; AAE17567.
XX
PT Composition comprising MTB39 antigen and MTB32A antigen from
PT Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS Claim 67; Page 96-97; 136pp; English.
XX
CC The present invention relates to fusion proteins containing at least two
CC Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected with
CC Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is a cDNA encoding
CC Mycobacterium species Ra32FL mature protein mutant, Ra32FLMutSA
XX
SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 6; Length 1002;
Best Local Similarity 100.0%; Pred. No. 5.6e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993
```

RESULT 10
AAX34251
ID AAX34251 standard; DNA; 1068 BP.
XX
AC AAX34251;
XX
DT 06-JUL-1999 (first entry)
XX
DE Mycobacterium species nucleic acid sequence 50D.
XX
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW hybridisation; detection; vaccine; immunisation; infection; ss.
XX
OS Mycobacterium sp.
XX
PN WO9909186-A2.
XX
PD 25-FEB-1999.
XX
PF 14-AUG-1998; 98WO-FR001813.
XX
PR 14-AUG-1997; 97FR-00010404.
PR 11-SEP-1997; 97FR-00011325.
XX
PA (INSP) INST PASTEUR.
XX
PI Gicquel B, Portnoie D, Lim E, Pelicic V, Guigueno A;
PI Goguet De La Salmoniere Y;
XX
DR WPI; 1999-181045/15.
DR P-PSDB; AAY05000.
XX
PT Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in infection-associated
PT protein expression.
XX
PS Claim 22; Fig 50D; 309pp; French.
XX
CC Sequences AAX34001-X34252 represent nucleic acids encoding secreted
CC proteins from various Mycobacterium species microorganisms. The
CC nucleotide sequences can be used as primers and probes for methods for

```
CC   detecting and identifying mycobacteria, especially belonging to the M.
CC   tuberculosis complex. The encoded proteins can be used in vaccines for
CC   immunisation against a bacterial or viral infection
XX
SQ   Sequence 1068 BP; 169 A; 352 C; 375 G; 172 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 2;  Length 1068;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 729

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 789

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 849

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 909

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 969

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1029

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1030 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1065


RESULT 11
AAX34252
ID   AAX34252 standard; DNA; 1143 BP.
XX
AC   AAX34252;
XX
DT   06-JUL-1999  (first entry)
XX
DE   Mycobacterium species nucleic acid sequence 50F.
XX
KW   Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW   hybridisation; detection; vaccine; immunisation; infection; ss.
XX
OS   Mycobacterium sp.
XX
PN   WO9909186-A2.
XX
PD   25-FEB-1999.
XX
PF   14-AUG-1998;   98WO-FR001813.
XX
PR   14-AUG-1997;   97FR-00010404.
PR   11-SEP-1997;   97FR-00011325.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Gicquel B,  Portnoie D,  Lim E,  Pelicic V,  Guigueno A;
PI   Goguet De La Salmoniere Y;
XX
DR   WPI; 1999-181045/15.
DR   P-PSDB; AAY04830.
XX
PT   Mycobacterial DNA vectors containing reporter constructs - for
PT   identifying coding or promoter sequences involved in infection-associated
PT   protein expression.
XX
PS   Claim 22; Fig 50F; 309pp; French.
XX
CC   Sequences AAX34001-X34252 represent nucleic acids encoding secreted
CC   proteins from various Mycobacterium species microorganisms. The
CC   nucleotide sequences can be used as primers and probes for methods for
CC   detecting and identifying mycobacteria, especially belonging to the M.
CC   tuberculosis complex. The encoded proteins can be used in vaccines for
CC   immunisation against a bacterial or viral infection
XX
SQ   Sequence 1143 BP; 189 A; 373 C; 395 G; 186 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 2;  Length 1143;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 804

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 864

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 924

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 984

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1044

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1104

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1105 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1140


RESULT 12
AAL40771
ID   AAL40771 standard; DNA; 1742 BP.
XX
AC   AAL40771;
XX
DT   03-OCT-2002  (first entry)
XX
DE   Nucleotide sequence encoding Ra12-WT1 fusion protein.
XX
KW   Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW   vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW   immunogen; cytokine; gene; ds.
XX
OS   Unidentified.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .1740
FT                   /*tag=  a
FT                   /product= "Ra12-WT1 fusion protein"
XX
PN   WO200125401-A2.
XX
PD   12-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US027652.
```

XX
PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J;
XX
DR WPI; 2001-266299/27.
DR P-PSDB; AAO22140.
XX
PT Recombinant nucleic acid molecule for producing high yield expression of
PT desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS Disclosure; Fig 4; 39pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Ra12-WT1 fusion protein
XX
SQ Sequence 1742 BP; 372 A; 569 C; 490 G; 311 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 4; Length 1742;
Best Local Similarity 100.0%; Pred. No. 5.7e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 13
ADA26353
ID ADA26353 standard; DNA; 2181 BP.
XX
AC ADA26353;

XX
DT 20-NOV-2003 (first entry)
XX
DE Mycobacterium MTB32-MTB39F fusion protein encoding DNA.
XX
KW ds; gene; fusion polypeptide; MTB32A; MTB39; antigen; MTB32A; MTB39;
KW MTB85A; tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS Chimeric.
OS Mycobacterium sp.
XX
FH Key Location/Qualifiers
FT CDS 4. .2175
FT /*tag= a
FT /product= "MTB32-MTB39F fusion polypeptide"
XX
PN WO2003070187-A2.
XX
PD 28-AUG-2003.
XX
PF 18-FEB-2003; 2003WO-US004903.
XX
PR 15-FEB-2002; 2002US-0357351P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J, Reed S;
XX
DR WPI; 2003-697554/66.
DR P-PSDB; ADA26354.
XX
PT New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT and MTB39 antigens, with or without the MTB85A antigen, from a
PT Mycobacterium species, useful for diagnosing, preventing and/or treating
PT tuberculosis infection.
XX
PS Claim 1; Fig 1; 112pp; English.
XX
CC The invention relates to a novel nucleic acid encoding a fusion
CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC A polypeptide of the invention has tuberculostatic activity. A
CC polynucelotide of the invention may have a use in gene therapy, and as a
CC vaccine. The methods and compositions of the present invention are useful
CC for diagnosing, preventing and/or treating tuberculosis infection. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 2181 BP; 341 A; 693 C; 792 G; 355 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 8; Length 2181;
Best Local Similarity 100.0%; Pred. No. 5.8e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897
```

```
Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993
```

RESULT 14
AAD47084
ID AAD47084 standard; DNA; 2190 BP.
XX
AC AAD47084;
XX
DT 29-AUG-2003 (revised)
DT 27-JAN-2003 (first entry)
XX
DE Mycobacterium sp. MTB72FMutSA fusion protein encoding DNA.
XX
KW Vaccine; immunity; diagnostic agent; gene therapy; TbH9; antigen;
KW Ra35MutSA; Ra12; MTB72MutSA; chimeric; gene; ds.
XX
OS Mycobacterium sp.
OS Mycobacterium tuberculosis.
OS Chimeric.
XX

FH	Key	Location/Qualifiers
FT	CDS	1. .2190
FT		/*tag= a
FT		/product= "MTB72FMutSA fusion protein"

XX
PN WO200272792-A2.
XX
PD 19-SEP-2002.
XX
PF 13-MAR-2002; 2002WO-US008223.
XX
PR 13-MAR-2001; 2001US-0275837P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Brannon M, Guderian J;
XX
DR WPI; 2002-759844/82.
DR P-PSDB; AAE29709.
XX
PT New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT tuberculosis.
XX
PS Disclosure; Page 92-93; 155pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC polynucleotide sequence encoding an antigen or an antigenic fragment from
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC polypeptide or its fragment. The Leishmania polynucleotide is selected
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides, as
CC in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is a
CC DNA encoding MTB72F fusion protein. This fusion protein comprises
CC Ra35MutSA mutant protein and Ra12 protein from M. tuberculosis and TbH9
CC protein from Mycobacterium sp. (Updated on 29-AUG-2003 to standardise OS
CC field)
XX
SQ Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 6; Length 2190;
Best Local Similarity 100.0%; Pred. No. 5.8e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 141

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 201

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 321

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 381

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        382 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 417
```

RESULT 15
AAD28343
ID AAD28343 standard; DNA; 2190 BP.
XX
AC AAD28343;
XX
DT 22-APR-2002 (first entry)
XX
DE Mycobacterium species MTB72FMutSA fusion protein encoding DNA.
XX
KW Fusion protein; antigen; serological sensitivity; immune response;
KW tuberculosis; infection; vaccine; MTB72FMutSA; Ra12-TbH9-Ra35MutSA;
KW mutant; mutein; ds.
XX
OS Mycobacterium sp.
XX

FH	Key	Location/Qualifiers
FT	CDS	1. .2190
FT		/*tag= a
FT		/product= "MTB72FMutSA fusion protein"
FT	misc_feature	22. .417
FT		/*tag= b
FT		/note= "Ra12 DNA fragment"
FT	misc_feature	424. .1596
FT		/*tag= c
FT		/note= "TbH9FL DNA fragment"
FT	misc_feature	1603. .2187
FT		/*tag= d
FT		/note= "Ra35 DNA fragment"
FT	mutation	replace(2128, T)
FT		/*tag= e

XX
PN WO200198460-A2.
XX
PD 27-DEC-2001.
XX
PF 20-JUN-2001; 2001WO-US019959.
XX
PR 20-JUN-2000; 2000US-00597796.
PR 01-FEB-2001; 2001US-0265737P.

```
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Reed S,  Alderson M;
XX
DR   WPI; 2002-147798/19.
DR   P-PSDB; AAE17573.
XX
PT   Composition comprising MTB39 antigen and MTB32A antigen from
PT   Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS   Claim 81; Page 108-109; 136pp; English.
XX
CC   The present invention relates to fusion proteins containing at least two
CC   Mycobacterium species antigens, nucleotides encoding them and
CC   compositions comprising such fusion proteins. The present invention
CC   particularly relates to nucleic acids encoding fusion proteins that
CC   include two or more individual M. tuberculosis antigens which increase
CC   the serological sensitivity of sera from individuals infected with
CC   tuberculosis and methods for their use in diagnosis, prevention and
CC   treatment of tuberculosis infection. Sequences of the invention are
CC   useful for eliciting an immune response in a mammal, e.g., human,
CC   immunised with BCG. They are useful in the diagnosis, treatment and
CC   prevention of Mycobacterium infection. The fusion proteins and the
CC   polynucleotides are useful as diagnostic tools in patients infected with
CC   Mycobacterium, in vitro and in vivo assays for detecting humoral
CC   antibodies or cell-mediated immunity against M. tuberculosis, for the
CC   diagnosis of an infection or monitoring of disease progression, as
CC   immunogens to generate or elicit a protective immune response in a
CC   patient and for raising anti-M. tuberculosis antibodies in a non-human
CC   animal. Sequences of the invention are also used as vaccines. MTB32A
CC   fusion proteins of the invention are useful as in vivo diagnostic agents
CC   for intradermal skin test. The present sequence is a DNA encoding
CC   Mycobacterium species MTB72FMutSA (Ra12-TbH9-Ra35MutSA) mutant fusion
CC   protein
XX
SQ   Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 6;  Length 2190;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 141

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 201

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 321

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 381

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        382 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 417


RESULT 16
AAL40773
ID   AAL40773 standard; DNA; 2191 BP.
XX
AC   AAL40773;
XX
DT   03-OCT-2002  (first entry)
XX
DE   Nucleotide sequence encoding Ra12-H9-32A fusion protein.
XX
KW   Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW   vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW   immunogen; cytokine; gene; ds.
XX
OS   Unidentified.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2190
FT                   /*tag=  a
FT                   /product= "Ra12-H9-32A fusion protein"
XX
PN   WO200125401-A2.
XX
PD   12-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US027652.
XX
PR   07-OCT-1999;   99US-0158585P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J;
XX
DR   WPI; 2001-266299/27.
DR   P-PSDB; AAO22142.
XX
PT   Recombinant nucleic acid molecule for producing high yield expression of
PT   desired fusion polypeptides, encodes fusion polypeptide comprising
PT   Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS   Disclosure; Fig 6; 39pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC   kDa C-terminal fragment of serine protease antigen MTB32A of
CC   Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC   The recombinant fusion nucleic acids and polypeptides are useful for
CC   providing stable and high yield expression of fusion polypeptides of both
CC   eukaryotic and prokaryotic origin and to encode a protein product for use
CC   as an antigen for detecting serum antibodies. The presence of serum
CC   antibodies to M. tuberculosis antigens in an individual indicates that
CC   the individual is infected with it. The fusion polypeptides are useful as
CC   sources of proteins for monitoring binding of serum antibodies to fusion
CC   proteins and as an immunogen to induce and/or enhance immune responses.
CC   The coding sequences can be ligated with a coding sequence of another
CC   molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC   can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC   represents the DNA encoding the Ra12-H9-32A fusion protein
XX
SQ   Sequence 2191 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 1 Other;

  Query Match             100.0%;  Score 396;  DB 4;  Length 2191;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 141

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 201

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 321

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 381

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
                ||||||||||||||||||||||||||||||||||||
Db        382 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 417
```

RESULT 17
ABK14128
ID ABK14128 standard; DNA; 2286 BP.
XX
AC ABK14128;
XX
DT 29-AUG-2003 (revised)
DT 08-MAY-2002 (first entry)
XX
DE DNA encoding antigenic fusion protein Ra12-TbH9-Ra35 (Mtb32-Mtb39).
XX
KW Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;
KW tuberculostatic; immunogen; vaccine; Mtb32-Mtb39; Ra12; TbH9; Ra35;
KW Ra12-TbH9-Ra35.
XX
OS Mycobacterium tuberculosis.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 42. .2231
FT /*tag= a
FT /product= "Mtb32-Mtb39"
FT /transl_except= (pos:498. .506, aa:Asn-Ala)
FT /transl_except= (pos:597. .605, aa:Ala-Gln)
FT /transl_except= (pos:798. .802, aa:Ala)
FT /note= "This codon has an apparent 2 nucleotide insertion
FT which alters the reading frame"
XX
PN US2002009459-A1.
XX
PD 24-JAN-2002.
XX
PF 07-APR-1999; 99US-00287849.
XX
PR 13-MAR-1997; 97US-00818112.
PR 01-OCT-1997; 97US-00942578.
PR 18-FEB-1998; 98US-00025197.
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (REED/) REED S G.
PA (SKEI/) SKEIKY Y A.
PA (DILL/) DILLON D C.
PA (ALDE/) ALDERSON M.
PA (CAMP/) CAMPOS-NETO A.
XX
PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;
XX
DR WPI; 2002-171134/22.
DR P-PSDB; AAU74588.
XX
PT New fusion proteins of Mycobacterium tuberculosis antigens, useful for
PT diagnosing, treating or preventing M. tuberculosis infection,
PT particularly as vaccine for treating or preventing tuberculosis.
XX
PS Example; Fig 1; 62pp; English.
XX
CC The invention relates to a purified polypeptide which induces an immune
CC response of Mycobacterium tuberculosis. Polypeptides of the invention are
CC useful for diagnosing, treating or preventing M. tuberculosis infection,
CC particularly tuberculosis infection. In particular, the polypeptides are
CC useful as a vaccine formulation with an adjuvant to afford long-term
CC protection in animals against the development of tuberculosis. The
CC protein coding sequence may be used to encode a protein product for use
CC as an immunogen to induce and/or enhance an immune response to M.
CC tuberculosis. This sequence represents DNA encoding an M. tuberculosis
CC fusion protein of the invention. (Updated on 29-AUG-2003 to standardise
CC OS field)
XX
SQ Sequence 2286 BP; 373 A; 719 C; 807 G; 381 T; 0 U; 6 Other;

Query Match 100.0%; Score 396; DB 6; Length 2286;
Best Local Similarity 100.0%; Pred. No. 5.8e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
                ||||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

RESULT 18
AAZ20194
ID AAZ20194 standard; DNA; 2287 BP.
XX
AC AAZ20194;
XX
DT 17-JAN-2000 (first entry)
XX
DE Mycobacterium tuberculosis antigen fusion protein Mtb32A DNA.
XX
KW Tuberculosis; antigen; fusion protein; Mtb32A; Ra12; TbH9; Ra35;
KW diagnosis; therapy; vaccine; immunogen; ss.
XX
OS Mycobacterium tuberculosis.
XX
FH Key Location/Qualifiers
FT CDS 42. .2231
FT /*tag= a
XX
PN WO9951748-A2.
XX
PD 14-OCT-1999.

XX
PF 07-APR-1999; 99WO-US007717.
XX
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Alderson M, Campos-Neto A;
XX
DR WPI; 1999-601610/51.
DR P-PSDB; AAY32059.
XX
PT New fusion proteins useful for diagnosis, prevention and treatment of
PT tuberculosis.
XX
PS Example; Fig 1A-B; 83pp; English.
XX
CC This DNA sequence includes a coding region for a recombinant
CC Mycobacterium tuberculosis tri-antigen fusion protein (see AAY32059),
CC termed Mtb32A, composed of the antigens Ra12, TbH9 and Ra35. The DNA is
CC useful for the recombinant production of the fusion protein. Coding
CC sequences for the antigens were modified by PCR in order to facilitate
CC their fusion and subsequent expression of the fusion protein. 3 Coding
CC sequences for Ra12, TbH9 and Ra25 were ligated to encode Mtb32A. The
CC invention provides fusion proteins (see AAY32059-71) containing at least
CC 2 M. tuberculosis antigens. The new fusion proteins and polynucleotides
CC encoding them are useful as vaccines for preventing tuberculosis
CC (claimed), for diagnosis (via in vitro assays or intradermal skin tests
CC for detection of anti-M. tuberculosis antibodies), monitoring of disease
CC progression, and treatment of tuberculosis. They are more effective
CC immunogens than mixtures of the individual protein components
XX
SQ Sequence 2287 BP; 374 A; 723 C; 805 G; 382 T; 0 U; 3 Other;

Query Match 100.0%; Score 396; DB 2; Length 2287;
Best Local Similarity 100.0%; Pred. No. 5.8e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

RESULT 19
AAD47083
ID AAD47083 standard; DNA; 2287 BP.
XX
AC AAD47083;
XX
DT 29-AUG-2003 (revised)
DT 27-JAN-2003 (first entry)
XX
DE Mycobacterium sp. MTB72F fusion protein encoding DNA.
XX
KW Vaccine; immunity; diagnostic agent; gene therapy; TbH9; antigen; Ra35;
KW Ra12; MTB72F; chimeric; gene; ds.
XX
OS Mycobacterium sp.
OS Mycobacterium tuberculosis.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 42. .2231
FT /*tag= a
FT /product= "MTB72F fusion protein"
XX
PN WO200272792-A2.
XX
PD 19-SEP-2002.
XX
PF 13-MAR-2002; 2002WO-US008223.
XX
PR 13-MAR-2001; 2001US-0275837P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Brannon M, Guderian J;
XX
DR WPI; 2002-759844/82.
DR P-PSDB; AAE29708.
XX
PT New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT tuberculosis.
XX
PS Disclosure; Page 87-90; 155pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC polynucleotide sequence encoding an antigen or an antigenic fragment from
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC polypeptide or its fragment. The Leishmania polynucleotide is selected
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides, as
CC in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is a
CC DNA encoding MTB72F fusion protein. This fusion protein comprises Ra12
CC and Ra35 protein from Mycobacterium tuberculosis and TbH9 protein from
CC Mycobacterium sp. (Updated on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 2287 BP; 372 A; 717 C; 814 G; 381 T; 0 U; 3 Other;

Query Match 100.0%; Score 396; DB 6; Length 2287;
Best Local Similarity 100.0%; Pred. No. 5.8e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242
```

```
Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

RESULT 20
AAD28342
ID AAD28342 standard; DNA; 2287 BP.
XX
AC AAD28342;
XX
DT 22-APR-2002 (first entry)
XX
DE Mycobacterium species MTB72F fusion protein encoding DNA.
XX
KW Fusion protein; antigen; serological sensitivity; immune response;
KW tuberculosis; infection; vaccine; MTB72F; Ra12-TbH9-Ra35; ds.
XX
OS Mycobacterium sp.
XX

FH	Key	Location/Qualifiers
FT	CDS	42. .2231
FT		/*tag= a
FT		/product= "MTB72F fusion protein"
FT	misc_feature	63. .458
FT		/*tag= b
FT		/note= "Ra12 DNA fragment"
FT	misc_feature	465. .1637
FT		/*tag= c
FT		/note= "TbH9FL DNA fragment"
FT	misc_feature	1644. .2228
FT		/*tag= d
FT		/note= "Ra35 DNA fragment"

XX
PN WO200198460-A2.
XX
PD 27-DEC-2001.
XX
PF 20-JUN-2001; 2001WO-US019959.
XX
PR 20-JUN-2000; 2000US-00597796.
PR 01-FEB-2001; 2001US-0265737P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Reed S, Alderson M;
XX
DR WPI; 2002-147798/19.
DR P-PSDB; AAE17572.
XX
PT Composition comprising MTB39 antigen and MTB32A antigen from
PT Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS Claim 62; Page 103-106; 136pp; English.
XX
CC The present invention relates to fusion proteins containing at least two
CC Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected with
CC Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is a DNA encoding
CC Mycobacterium species MTB72F (Ra12-TbH9-Ra35) fusion protein
XX
SQ Sequence 2287 BP; 372 A; 717 C; 814 G; 381 T; 0 U; 3 Other;

Query Match 100.0%; Score 396; DB 6; Length 2287;
Best Local Similarity 100.0%; Pred. No. 5.8e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

RESULT 21
ADA26360
ID ADA26360 standard; DNA; 2451 BP.
XX
AC ADA26360;
XX
DT 20-NOV-2003 (first entry)
XX
DE Mycobacterium MTB72F-DPV (fusion MTB81F) protein encoding DNA.
XX
KW ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS Chimeric.
OS Mycobacterium sp.
XX

FH	Key	Location/Qualifiers
FT	CDS	4. .2445
FT		/*tag= a
FT		/product= " MTB72F-DPV (fusion MTB81F) protein"

XX

PN WO2003070187-A2.
XX
PD 28-AUG-2003.
XX
PF 18-FEB-2003; 2003WO-US004903.
XX
PR 15-FEB-2002; 2002US-0357351P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J, Reed S;
XX
DR WPI; 2003-697554/66.
DR P-PSDB; ADA26367.
XX
PT New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT and MTB39 antigens, with or without the MTB85A antigen, from a
PT Mycobacterium species, useful for diagnosing, preventing and/or treating
PT tuberculosis infection.
XX
PS Claim 84; Fig 8; 112pp; English.
XX
CC The invention relates to a novel nucleic acid encoding a fusion
CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC A polypeptide of the invention has tuberculostatic activity. A
CC polynucelotide of the invention may have a use in gene therapy, and as a
CC vaccine. The methods and compositions of the present invention are useful
CC for diagnosing, preventing and/or treating tuberculosis infection. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 2451 BP; 389 A; 785 C; 870 G; 407 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 8; Length 2451;
Best Local Similarity 100.0%; Pred. No. 5.9e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 22
ADA26359
ID ADA26359 standard; DNA; 2487 BP.
XX
AC ADA26359;
XX
DT 20-NOV-2003 (first entry)
XX
DE Mycobacterium MTB72F-MTI (fusion MTB83F) protein encoding DNA.
XX
KW ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS Chimeric.
OS Mycobacterium sp.
XX
FH Key Location/Qualifiers
FT CDS 4. .2481
FT /*tag= a
FT /product= "MTB72F-MTI (fusion MTB83F) protein"
XX
PN WO2003070187-A2.
XX
PD 28-AUG-2003.
XX
PF 18-FEB-2003; 2003WO-US004903.
XX
PR 15-FEB-2002; 2002US-0357351P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J, Reed S;
XX
DR WPI; 2003-697554/66.
DR P-PSDB; ADA26366.
XX
PT New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT and MTB39 antigens, with or without the MTB85A antigen, from a
PT Mycobacterium species, useful for diagnosing, preventing and/or treating
PT tuberculosis infection.
XX
PS Claim 84; Fig 7; 112pp; English.
XX
CC The invention relates to a novel nucleic acid encoding a fusion
CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC A polypeptide of the invention has tuberculostatic activity. A
CC polynucelotide of the invention may have a use in gene therapy, and as a
CC vaccine. The methods and compositions of the present invention are useful
CC for diagnosing, preventing and/or treating tuberculosis infection. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 2487 BP; 394 A; 784 C; 896 G; 413 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 8; Length 2487;
Best Local Similarity 100.0%; Pred. No. 5.9e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 23
ADA26358
ID   ADA26358 standard; DNA; 2637 BP.
XX
AC   ADA26358;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB72F-Erd14 (fusion MTB89F) protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2631
FT                   /*tag=  a
FT                   /product= "MTB72F-Erd14 (fusion MTB89F)"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26365.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 84; Fig 6; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 2637 BP; 428 A; 840 C; 928 G; 441 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 8;  Length 2637;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 24
AAD47110
ID   AAD47110 standard; DNA; 2808 BP.
XX
AC   AAD47110;
XX
DT   29-AUG-2003  (revised)
DT   27-JAN-2003  (first entry)
XX
DE   Mycobacterium sp. MTB72F-Leishmania sp. MAPS (aka r95f)-fusion DNA.
XX
KW   Vaccine; immunity; diagnostic agent; gene therapy; MTB72F; MAPS;
KW   chimeric; gene; ds.
XX
OS   Mycobacterium sp.
OS   Leishmania sp.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2796
FT                   /*tag=  a
FT                   /product= "Mycobacterium sp. MTB72F-Leishmania sp. MAPS
FT                   (aka r95f) fusion protein"
XX
PN   WO200272792-A2.
XX
PD   19-SEP-2002.
XX
PF   13-MAR-2002; 2002WO-US008223.
XX
PR   13-MAR-2001; 2001US-0275837P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Brannon M,  Guderian J;
XX
DR   WPI; 2002-759844/82.
DR   P-PSDB; AAE29731.
XX
PT   New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT   M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT   against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT   tuberculosis.
XX
PS   Example 6; Page 128-129; 155pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC   polynucleotide sequence encoding an antigen or an antigenic fragment from
CC   Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC   polypeptide or its fragment. The Leishmania polynucleotide is selected
CC   from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
```

CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides, as
CC in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is
CC Mycobacterium sp. MTB72F-Leishmania sp. thiol specific antioxidant [TSA;
CC MAPS (aka r95f)] fusion DNA. This sequence comprises Mycobacterium sp.
CC MTB72F (a 72 kDa poly-protein fusion construct comprising Ra12-TbH9-Ra35)
CC linked to the Leishmania TSA (MAPS). (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
SQ Sequence 2808 BP; 466 A; 889 C; 982 G; 471 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 6; Length 2808;
Best Local Similarity 100.0%; Pred. No. 5.9e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 25
ADA26357
ID ADA26357 standard; DNA; 2808 BP.
XX
AC ADA26357;
XX
DT 20-NOV-2003 (first entry)
XX
DE Mycobacterium MTB72F-MAPS (fusion r95F) protein encoding DNA.
XX
KW ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS Chimeric.
OS Mycobacterium sp.
XX
FH Key Location/Qualifiers
FT CDS 4. .2796
FT /*tag= a
FT /product= "MTB72F-MAPS (fusion r95F)"
XX
PN WO2003070187-A2.
XX
PD 28-AUG-2003.
XX
PF 18-FEB-2003; 2003WO-US004903.
XX
PR 15-FEB-2002; 2002US-0357351P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J, Reed S;
XX
DR WPI; 2003-697554/66.
DR P-PSDB; ADA26364.
XX
PT New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT and MTB39 antigens, with or without the MTB85A antigen, from a
PT Mycobacterium species, useful for diagnosing, preventing and/or treating
PT tuberculosis infection.
XX
PS Disclosure; Fig 5; 112pp; English.
XX
CC The invention relates to a novel nucleic acid encoding a fusion
CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC A polypeptide of the invention has tuberculostatic activity. A
CC polynucelotide of the invention may have a use in gene therapy, and as a
CC vaccine. The methods and compositions of the present invention are useful
CC for diagnosing, preventing and/or treating tuberculosis infection. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 2808 BP; 466 A; 889 C; 982 G; 471 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 8; Length 2808;
Best Local Similarity 100.0%; Pred. No. 5.9e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 26
ADA26355
ID ADA26355 standard; DNA; 3030 BP.
XX
AC ADA26355;
XX
DT 20-NOV-2003 (first entry)
XX
DE Mycobacterium MTB-102F fusion protein encoding DNA.
XX
KW ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW tuberculosis; tuberculostatic; gene therapy; vaccine.

```
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .3030
FT                   /*tag= a
FT                   /product= "MTB-120F fusion protein"
FT                   /note= "No stop codon given"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26356.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 5; Fig 3; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 3030 BP; 498 A; 966 C; 1062 G; 504 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 8;  Length 3030;
  Best Local Similarity   100.0%;  Pred. No. 6e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 654

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 714

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 774

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 834

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 894

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 954

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        955 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 990


RESULT 27
ADA26363
ID   ADA26363 standard; DNA; 3060 BP.
XX
AC   ADA26363;
XX
DT   20-NOV-2003  (first entry)
XX
DE   M. bovis MTB72F and 85b complex antigen (fusion MTB103F) DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine; MTB72F; 85B.
XX
OS   Chimeric.
OS   Mycobacterium bovis.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .3054
FT                   /*tag= a
FT                   /product= "MTB72F and 85b complex (fusion MTB103F)"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26370.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 84; Fig 11; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 3060 BP; 503 A; 977 C; 1070 G; 510 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 8;  Length 3060;
  Best Local Similarity   100.0%;  Pred. No. 6e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204
```

```
Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 28
ADA26362
ID ADA26362 standard; DNA; 3104 BP.
XX
AC ADA26362;
XX
DT 20-NOV-2003 (first entry)
XX
DE Mycobacterium MTB72F-hTCC#1 (fusion MTB102tm2F) protein encoding DNA.
XX
KW ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS Chimeric.
OS Mycobacterium sp.
XX
FH Key Location/Qualifiers
FT CDS 4. .3072
FT /*tag= a
FT /product= "MTB72F-hTCC#1 (fusion MTB102tm2F) protein"
XX
PN WO2003070187-A2.
XX
PD 28-AUG-2003.
XX
PF 18-FEB-2003; 2003WO-US004903.
XX
PR 15-FEB-2002; 2002US-0357351P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J, Reed S;
XX
DR WPI; 2003-697554/66.
DR P-PSDB; ADA26369.
XX
PT New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT and MTB39 antigens, with or without the MTB85A antigen, from a
PT Mycobacterium species, useful for diagnosing, preventing and/or treating
PT tuberculosis infection.
XX
PS Claim 84; Fig 10; 112pp; English.
XX
CC The invention relates to a novel nucleic acid encoding a fusion
CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC A polypeptide of the invention has tuberculostatic activity. A
CC polynucelotide of the invention may have a use in gene therapy, and as a
CC vaccine. The methods and compositions of the present invention are useful
CC for diagnosing, preventing and/or treating tuberculosis infection. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 3104 BP; 510 A; 968 C; 1104 G; 522 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 8; Length 3104;
Best Local Similarity 100.0%; Pred. No. 6e-84;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 29
ADA26361
ID ADA26361 standard; DNA; 3474 BP.
XX
AC ADA26361;
XX
DT 20-NOV-2003 (first entry)
XX
DE Mycobacterium MTB72F-mTCC#2 (fusion MTB114F) protein encoding DNA.
XX
KW ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS Chimeric.
OS Mycobacterium sp.
XX
FH Key Location/Qualifiers
FT CDS 4. .3468
FT /*tag= a
FT /product= "MTB72F-mTCC#2 (fusion MTB114F) protein"
XX
PN WO2003070187-A2.
XX
PD 28-AUG-2003.
XX
PF 18-FEB-2003; 2003WO-US004903.
XX
PR 15-FEB-2002; 2002US-0357351P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J, Reed S;
XX
DR WPI; 2003-697554/66.
DR P-PSDB; ADA26368.
XX
PT New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT and MTB39 antigens, with or without the MTB85A antigen, from a
PT Mycobacterium species, useful for diagnosing, preventing and/or treating
PT tuberculosis infection.
XX
PS Claim 84; Fig 9; 112pp; English.

```
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 3474 BP; 548 A; 1131 C; 1252 G; 543 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 396;  DB 8;  Length 3474;
  Best Local Similarity   100.0%;  Pred. No. 6e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 30
AAI99682_01
Continuation (2 of 45) of AAI99682 from base 100001 (Mycobacterium tuberculosis strain H
WP Sequence split into 45 fragments  LOCUS AAI99682 Accession Aai99682
WP      Fragment Name       Begin      End
WP      AAI99682_00             1    110000
WP      AAI99682_01        100001    210000
WP      AAI99682_02        200001    310000
WP      AAI99682_03        300001    410000
WP      AAI99682_04        400001    510000
WP      AAI99682_05        500001    610000
WP      AAI99682_06        600001    710000
WP      AAI99682_07        700001    810000
WP      AAI99682_08        800001    910000
WP      AAI99682_09        900001   1010000
WP      AAI99682_10       1000001   1110000
WP      AAI99682_11       1100001   1210000
WP      AAI99682_12       1200001   1310000
WP      AAI99682_13       1300001   1410000
WP      AAI99682_14       1400001   1510000
WP      AAI99682_15       1500001   1610000
WP      AAI99682_16       1600001   1710000
WP      AAI99682_17       1700001   1810000
WP      AAI99682_18       1800001   1910000
WP      AAI99682_19       1900001   2010000
WP      AAI99682_20       2000001   2110000
WP      AAI99682_21       2100001   2210000
WP      AAI99682_22       2200001   2310000
WP      AAI99682_23       2300001   2410000
WP      AAI99682_24       2400001   2510000
WP      AAI99682_25       2500001   2610000
WP      AAI99682_26       2600001   2710000
WP      AAI99682_27       2700001   2810000
WP      AAI99682_28       2800001   2910000
WP      AAI99682_29       2900001   3010000
WP      AAI99682_30       3000001   3110000
WP      AAI99682_31       3100001   3210000
WP      AAI99682_32       3200001   3310000
WP      AAI99682_33       3300001   3410000
WP      AAI99682_34       3400001   3510000
WP      AAI99682_35       3500001   3610000
WP      AAI99682_36       3600001   3710000
WP      AAI99682_37       3700001   3810000
WP      AAI99682_38       3800001   3910000
WP      AAI99682_39       3900001   4010000
WP      AAI99682_40       4000001   4110000
WP      AAI99682_41       4100001   4210000
WP      AAI99682_42       4200001   4310000
WP      AAI99682_43       4300001   4410000
WP      AAI99682_44       4400001   4411529

  Query Match             100.0%;  Score 396;  DB 4;  Length 110000;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-84;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51815 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 51874

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51875 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 51934

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51935 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 51994

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51995 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 52054

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52055 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 52114

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52115 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 52174

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db      52175 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 52210


Search completed: April 30, 2004, 05:15:25
Job time : 297 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 30, 2004, 05:10:50 ; Search time 293 Seconds
(without alignments)
6113.624 Million cell updates/sec

Title: US-09-684-215B-3
Perfect score: 396
Sequence: 1 acggccgcgtccgataactt.....tggccgagggacccccggcc 396

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	396	100.0	702	9	US-09-287-849-27	Sequence 27, Appl
2	396	100.0	702	15	US-10-359-460-27	Sequence 27, Appl
3	396	100.0	1002	15	US-10-098-732A-3	Sequence 3, Appli
4	396	100.0	1002	15	US-10-098-732A-5	Sequence 5, Appli
5	396	100.0	1068	9	US-09-712-363-15	Sequence 15, Appl
6	396	100.0	2181	16	US-10-369-983-1	Sequence 1, Appli
7	396	100.0	2190	15	US-10-098-732A-17	Sequence 17, Appl
8	396	100.0	2287	9	US-09-287-849-1	Sequence 1, Appli
9	396	100.0	2287	15	US-10-359-460-1	Sequence 1, Appli
10	396	100.0	2287	15	US-10-098-732A-15	Sequence 15, Appl
11	396	100.0	2287	16	US-10-359-459-1	Sequence 1, Appli
12	396	100.0	2451	16	US-10-369-983-8	Sequence 8, Appli
13	396	100.0	2487	16	US-10-369-983-7	Sequence 7, Appli
14	396	100.0	2637	16	US-10-369-983-6	Sequence 6, Appli
15	396	100.0	2808	15	US-10-098-732A-64	Sequence 64, Appl
16	396	100.0	2808	16	US-10-369-983-5	Sequence 5, Appli
17	396	100.0	3030	16	US-10-369-983-3	Sequence 3, Appli
18	396	100.0	3060	16	US-10-369-983-11	Sequence 11, Appl
19	396	100.0	3104	16	US-10-369-983-10	Sequence 10, Appl
20	396	100.0	3474	16	US-10-369-983-9	Sequence 9, Appli
21	394.4	99.6	447	15	US-10-193-002-4	Sequence 4, Appli
22	394.4	99.6	447	15	US-10-084-843-4	Sequence 4, Appli
23	394.4	99.6	447	15	US-10-098-732A-9	Sequence 9, Appli
24	391.2	98.8	1872	15	US-10-193-002-17	Sequence 17, Appl
25	391.2	98.8	1872	15	US-10-084-843-17	Sequence 17, Appl
26	391.2	98.8	1872	15	US-10-098-732A-1	Sequence 1, Appli
27	354	89.4	675	9	US-09-759-143-822	Sequence 822, App
28	354	89.4	675	9	US-09-780-669-822	Sequence 822, App
29	354	89.4	675	9	US-09-822-827-822	Sequence 822, App
30	354	89.4	675	9	US-09-895-793-822	Sequence 822, App
31	354	89.4	675	9	US-09-895-814-822	Sequence 822, App
32	354	89.4	675	14	US-10-012-896-822	Sequence 822, App
33	354	89.4	675	15	US-10-144-678A-822	Sequence 822, App
34	354	89.4	675	15	US-10-294-025-822	Sequence 822, App
35	354	89.4	822	9	US-09-736-457-1862	Sequence 1862, Ap
36	354	89.4	822	9	US-09-902-941-1862	Sequence 1862, Ap
37	354	89.4	822	9	US-09-849-626-1862	Sequence 1862, Ap
38	354	89.4	822	13	US-10-283-017-1862	Sequence 1862, Ap
39	354	89.4	822	15	US-10-017-754-1862	Sequence 1862, Ap
40	354	89.4	822	15	US-10-113-872-1862	Sequence 1862, Ap
41	354	89.4	894	13	US-09-878-722-236	Sequence 236, App
42	354	89.4	894	13	US-09-904-456-236	Sequence 236, App
43	354	89.4	900	9	US-09-735-705-353	Sequence 353, App
44	354	89.4	900	9	US-09-850-716A-353	Sequence 353, App
45	354	89.4	900	9	US-09-897-778-353	Sequence 353, App

ALIGNMENTS

RESULT 1
US-09-287-849-27
; Sequence 27, Application US/09287849
; Patent No. US20020009459A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1
; NAME/KEY: CDS

; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-09-287-849-27

Query Match 100.0%; Score 396; DB 9; Length 702;
Best Local Similarity 100.0%; Pred. No. 3e-107;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 2
US-10-359-460-27
; Sequence 27, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-10-359-460-27

Query Match 100.0%; Score 396; DB 15; Length 702;
Best Local Similarity 100.0%; Pred. No. 3e-107;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 3
US-10-098-732A-3
; Sequence 3, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderian, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13

```
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 1002
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Ra35 mature
US-10-098-732A-3

  Query Match            100.0%;  Score 396;  DB 15;  Length 1002;
  Best Local Similarity  100.0%;  Pred. No. 3e-107;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993


RESULT 4
US-10-098-732A-5
; Sequence 5, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1002
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Ra35FLMutSA
US-10-098-732A-5

  Query Match            100.0%;  Score 396;  DB 15;  Length 1002;
  Best Local Similarity  100.0%;  Pred. No. 3e-107;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993


RESULT 5
US-09-712-363-15
; Sequence 15, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
;  APPLICANT: Eisenberg, David
;  APPLICANT:  Rotstein, Sergio H.
;  APPLICANT:  Marcotte, Edward M.
;  TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
;  TITLE OF INVENTION:  INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
;  FILE REFERENCE: 07419-032001
;  CURRENT APPLICATION NUMBER: US/09/712,363
;  CURRENT FILING DATE:  2000-11-13
;  PRIOR APPLICATION NUMBER: PCT/US00/02246
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: 60/179,531
;  PRIOR FILING DATE: 2000-02-01
;  PRIOR APPLICATION NUMBER: 60/117,844
;  PRIOR FILING DATE: 1999-01-29
;  PRIOR APPLICATION NUMBER: 60/118,206,
;  PRIOR FILING DATE: 1999-02-01
;  PRIOR APPLICATION NUMBER: 60/126,593
;  PRIOR FILING DATE: 1999-03-26
;  PRIOR APPLICATION NUMBER: 60/134,093
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/134,092
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/165,124
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/165,086
;  PRIOR FILING DATE: 1999-11-12
;  NUMBER OF SEQ ID NOS: 292
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 1068
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
US-09-712-363-15

  Query Match            100.0%;  Score 396;  DB 9;  Length 1068;
  Best Local Similarity  100.0%;  Pred. No. 3e-107;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
```

```
Db        670 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 729

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120

Db        730 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 789

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180

Db        790 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 849

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240

Db        850 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 909

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300

Db        910 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 969

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360

Db        970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1029

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396

Db       1030 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1065
```

RESULT 6
US-10-369-983-1
; Sequence 1, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2181
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mutated
; OTHER INFORMATION: MTB32-MTB39F fusion protein (MTB32MutSA)
US-10-369-983-1

Query Match 100.0%; Score 396; DB 16; Length 2181;
Best Local Similarity 100.0%; Pred. No. 3e-107;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60

Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120

Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180

Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240

Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300

Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360

Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396

Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993
```

RESULT 7
US-10-098-732A-17
; Sequence 17, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderian, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB72FMutSA
; OTHER INFORMATION: (Ra12-TbH9-Ra35MutSA)
US-10-098-732A-17

Query Match 100.0%; Score 396; DB 15; Length 2190;
Best Local Similarity 100.0%; Pred. No. 3e-107;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60

Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120

Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 141

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180

Db        142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 201

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240

Db        202 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300

Db        262 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 321

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360

Db        322 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 381

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396

Db        382 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 417
```

```
RESULT 8
US-09-287-849-1
; Sequence 1, Application US/09287849
; Patent No. US20020009459A1
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Alderson, Mark
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION:  Fusion Protiens of Mycobacterium tuberculosis Antigens
;  TITLE OF INVENTION:  and Their Uses
;  FILE REFERENCE:  014058-009020US
;  CURRENT APPLICATION NUMBER: US/09/287,849
;  CURRENT FILING DATE:  1999-04-07
;  PRIOR APPLICATION NUMBER:  US 08/818,112
;  PRIOR FILING DATE:  1997-03-13
;  PRIOR APPLICATION NUMBER:  US 08/942,578
;  PRIOR FILING DATE:  1997-10-01
;  PRIOR APPLICATION NUMBER:  US 09/025,197
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 09/056,556
;  PRIOR FILING DATE:  1998-04-07
;  PRIOR APPLICATION NUMBER:  US 09/223,040
;  PRIOR FILING DATE:  1998-12-30
;  NUMBER OF SEQ ID NOS:  46
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2287
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
;   OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39
;   OTHER INFORMATION: fusion)
;   NAME/KEY: modified_base
;   LOCATION: (30)
;   OTHER INFORMATION: n = g, a, c or t
;   NAME/KEY: modified_base
;   LOCATION: (33)
;   OTHER INFORMATION: n = g, a, c or t
;   NAME/KEY: CDS
;   LOCATION: (42)..(2231)
;   NAME/KEY: modified_base
;   LOCATION: (2270)
;   OTHER INFORMATION: n = g, a, c or t
US-09-287-849-1

  Query Match             100.0%;  Score 396;  DB 9;  Length 2287;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-107;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458


RESULT 9
US-10-359-460-1
; Sequence 1, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Alderson, Mark
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
;  TITLE OF INVENTION:  and Their Uses
;  FILE REFERENCE: 014058-009020US
;  CURRENT APPLICATION NUMBER: US/10/359,460
;  CURRENT FILING DATE:  2003-02-05
;  PRIOR APPLICATION NUMBER: US/09/287,849
;  PRIOR FILING DATE: 1999-04-07
;  PRIOR APPLICATION NUMBER: US 08/818,112
;  PRIOR FILING DATE: 1997-03-13
;  PRIOR APPLICATION NUMBER: US 08/942,578
;  PRIOR FILING DATE: 1997-10-01
;  PRIOR APPLICATION NUMBER: US 09/025,197
;  PRIOR FILING DATE: 1998-02-18
;  PRIOR APPLICATION NUMBER: US 09/056,556
;  PRIOR FILING DATE: 1998-04-07
;  PRIOR APPLICATION NUMBER: US 09/223,040
;  PRIOR FILING DATE: 1998-12-30
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2287
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
;   OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39
;   OTHER INFORMATION: fusion)
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (30)
;   OTHER INFORMATION: n = g, a, c or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (33)
;   OTHER INFORMATION: n = g, a, c or t
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (42)..(2231)
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (2270)
;   OTHER INFORMATION: n = g, a, c or t
US-10-359-460-1

  Query Match             100.0%;  Score 396;  DB 15;  Length 2287;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-107;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122
```

```
Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

RESULT 10
US-10-098-732A-15
; Sequence 15, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderian, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
; OTHER INFORMATION: protein MTB72F (Ra12-TbH9-Ra35 or MTB32-MTB39
; OTHER INFORMATION: fusion)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
; OTHER INFORMATION: MTB72F
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(2287)
; OTHER INFORMATION: n = g, a, c or t
US-10-098-732A-15

Query Match 100.0%; Score 396; DB 15; Length 2287;
Best Local Similarity 100.0%; Pred. No. 3.1e-107;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

RESULT 11
US-10-359-459-1
; Sequence 1, Application US/10359459
; Publication No. US20040013677A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009010US
; CURRENT APPLICATION NUMBER: US/10/359,459
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
; OTHER INFORMATION: protein Ra12-TbH9-Ra35
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (30)
; OTHER INFORMATION: n = g, a, c or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (33)
; OTHER INFORMATION: n = g, a, c or t
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2270)
; OTHER INFORMATION: n = g, a, c or t
US-10-359-459-1

Query Match 100.0%; Score 396; DB 16; Length 2287;
Best Local Similarity 100.0%; Pred. No. 3.1e-107;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458


RESULT 12
US-10-369-983-8
; Sequence 8, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 2451
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB81F (MTB72F-DPV)
US-10-369-983-8

  Query Match             100.0%;  Score 396;  DB 16;  Length 2451;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-107;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 13
US-10-369-983-7
; Sequence 7, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 2487
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB83F (MTB72F-MTI)
US-10-369-983-7

  Query Match             100.0%;  Score 396;  DB 16;  Length 2487;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-107;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 14
US-10-369-983-6
; Sequence 6, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 2637
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB89F (MTB72F-Erd14)
US-10-369-983-6

  Query Match            100.0%;  Score 396;  DB 16;  Length 2637;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-107;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 15
US-10-098-732A-64
; Sequence 64, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
;   LENGTH: 2808
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:MTB72F-MAPS
;   OTHER INFORMATION: (r95f) fusion construct, TB MTB72F (Ra12-TbH9-Ra35)
;   OTHER INFORMATION: linked to Leishmania thiol-specific antioxidant
;   OTHER INFORMATION: (TSA or MAPS)
US-10-098-732A-64

  Query Match            100.0%;  Score 396;  DB 15;  Length 2808;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-107;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 16
US-10-369-983-5
; Sequence 5, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 2808
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: R95F (MTB72F-MAPS)
US-10-369-983-5

  Query Match            100.0%;  Score 396;  DB 16;  Length 2808;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-107;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 17
US-10-369-983-3
; Sequence 3, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3030
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB-102F fusion
; OTHER INFORMATION: protein
US-10-369-983-3

Query Match 100.0%; Score 396; DB 16; Length 3030;
Best Local Similarity 100.0%; Pred. No. 3.1e-107;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 654

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 714

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 774

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 834

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 894

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 954

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        955 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 990
```

RESULT 18
US-10-369-983-11
; Sequence 11, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 3060
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:fusion protein
; OTHER INFORMATION: MTB103F (MTB72F-85b)
US-10-369-983-11

Query Match 100.0%; Score 396; DB 16; Length 3060;
Best Local Similarity 100.0%; Pred. No. 3.1e-107;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

```
RESULT 19
US-10-369-983-10
; Sequence 10, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 3104
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB102tm2F (MTB102FTM, MTB72F-hTCC#1)
US-10-369-983-10

  Query Match             100.0%;  Score 396;  DB 16;  Length 3104;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-107;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 20
US-10-369-983-9
; Sequence 9, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 3474
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB114F (MTB72F-mTCC#2)
US-10-369-983-9

  Query Match             100.0%;  Score 396;  DB 16;  Length 3474;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-107;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 21
US-10-193-002-4
; Sequence 4, Application US/10193002
; Publication No. US20030135026A1
;  GENERAL INFORMATION:
;       APPLICANT: Reed, Steven G.
;                  Skeiky, Yasir A.W.
;                  Dillon, Davin C.
;                  Campos-Neto, Antonia
;                  Houghton, Raymond
;                  Vedvick, Thomas S.
;                  Twardzik, Daniel R.
;                  Lodes, Michael J.
;                  Hendrickson, Ronald C.
;       TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                           TUBERCULOSIS
;       NUMBER OF SEQUENCES: 350
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SEED and BERRY LLP
;            STREET: 6300 Columbia Center, 701 Fifth Avenue
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: USA
;            ZIP: 98104-7092
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
```

```
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/193,002
;             FILING DATE: 10-Jul-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/072,596
;             FILING DATE: 05-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Maki, David J.
;             REGISTRATION NUMBER: 31,392
;             REFERENCE/DOCKET NUMBER: 210121.417C9
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 447 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-193-002-4

  Query Match             99.6%;  Score 394.4;  DB 15;  Length 447;
  Best Local Similarity   99.7%;  Pred. No. 8.8e-107;
  Matches  395;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406


RESULT 22
US-10-084-843-4
; Sequence 4, Application US/10084843
; Publication No. US20030143243A1
;   GENERAL INFORMATION:
;        APPLICANT: Reed, Steven G.
;                   Skeiky, Yasir A.W.
;                   Dillon, Davin C.
;                   Campos-Neto, Antonio
;                   Houghton, Raymond
;                   Vedvick, Thomas S.
;                   Twardzik, Daniel R.
;                   Lodes, Michael J.
;                   Hendrickson, Ronald C.
;        TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;                            AND DIAGNOSIS OF TUBERCULOSIS
;        NUMBER OF SEQUENCES: 355
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SEED and BERRY LLP
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/084,843
;             FILING DATE: 25-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/072,967
;             FILING DATE: 05-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Maki, David J.
;             REGISTRATION NUMBER: 31,392
;             REFERENCE/DOCKET NUMBER: 210121.411C9
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 447 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-084-843-4

  Query Match             99.6%;  Score 394.4;  DB 15;  Length 447;
  Best Local Similarity   99.7%;  Pred. No. 8.8e-107;
  Matches  395;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406


RESULT 23
US-10-098-732A-9
; Sequence 9, Application US/10098732A
```

```
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 447
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:MTBRa12 (MTB32A
;   OTHER INFORMATION: C-terminus)
US-10-098-732A-9

  Query Match             99.6%;  Score 394.4;  DB 15;  Length 447;
  Best Local Similarity   99.7%;  Pred. No. 8.8e-107;
  Matches  395;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406


RESULT 24
US-10-193-002-17
; Sequence 17, Application US/10193002
; Publication No. US20030135026A1
;   GENERAL INFORMATION:
;        APPLICANT: Reed, Steven G.
;                   Skeiky, Yasir A.W.
;                   Dillon, Davin C.
;                   Campos-Neto, Antonia
;                   Houghton, Raymond
;                   Vedvick, Thomas S.
;                   Twardzik, Daniel R.
;                   Lodes, Michael J.
;                   Hendrickson, Ronald C.
;        TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                            TUBERCULOSIS
;        NUMBER OF SEQUENCES: 350
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SEED and BERRY LLP
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/193,002
;             FILING DATE: 10-Jul-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/072,596
;             FILING DATE: 05-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Maki, David J.
;             REGISTRATION NUMBER: 31,392
;             REFERENCE/DOCKET NUMBER: 210121.417C9
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 17:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1872 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-193-002-17

  Query Match             98.8%;  Score 391.2;  DB 15;  Length 1872;
  Best Local Similarity   99.2%;  Pred. No. 7.9e-106;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153


RESULT 25
US-10-084-843-17
; Sequence 17, Application US/10084843
; Publication No. US20030143243A1
```

```
;  GENERAL INFORMATION:
;       APPLICANT: Reed, Steven G.
;                  Skeiky, Yasir A.W.
;                  Dillon, Davin C.
;                  Campos-Neto, Antonio
;                  Houghton, Raymond
;                  Vedvick, Thomas S.
;                  Twardzik, Daniel R.
;                  Lodes, Michael J.
;                  Hendrickson, Ronald C.
;       TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;                           AND DIAGNOSIS OF TUBERCULOSIS
;       NUMBER OF SEQUENCES: 355
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SEED and BERRY LLP
;            STREET: 6300 Columbia Center, 701 Fifth Avenue
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: USA
;            ZIP: 98104-7092
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/084,843
;            FILING DATE: 25-Feb-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/072,967
;            FILING DATE: 05-MAY-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Maki, David J.
;            REGISTRATION NUMBER: 31,392
;            REFERENCE/DOCKET NUMBER: 210121.411C9
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (206) 622-4900
;            TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 17:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1872 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-084-843-17

  Query Match             98.8%;  Score 391.2;  DB 15;  Length 1872;
  Best Local Similarity   99.2%;  Pred. No. 7.9e-106;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153


RESULT 26
US-10-098-732A-1
; Sequence 1, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1872
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: MTB32A (Ra35FL)
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1)..(1872)
;   OTHER INFORMATION: n = g, a, c or t
US-10-098-732A-1

  Query Match             98.8%;  Score 391.2;  DB 15;  Length 1872;
  Best Local Similarity   99.2%;  Pred. No. 7.9e-106;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153


RESULT 27
US-09-759-143-822
```

```
; Sequence 822, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-143-822

  Query Match             89.4%;  Score 354;  DB 9;  Length 675;
  Best Local Similarity   95.7%;  Pred. No. 7.5e-95;
  Matches  379;  Conservative    0;  Mismatches    5;  Indels   12;  Gaps    1;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||           |||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC-----------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405


RESULT 28
US-09-780-669-822
; Sequence 822, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C24
;  CURRENT APPLICATION NUMBER: US/09/780,669
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS:  943
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-780-669-822

  Query Match             89.4%;  Score 354;  DB 9;  Length 675;
  Best Local Similarity   95.7%;  Pred. No. 7.5e-95;
  Matches  379;  Conservative    0;  Mismatches    5;  Indels   12;  Gaps    1;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||           |||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC-----------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405


RESULT 29
US-09-822-827-822
; Sequence 822, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C1
```

```
;  CURRENT APPLICATION NUMBER: US/09/822,827
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-822-827-822

  Query Match             89.4%;  Score 354;  DB 9;  Length 675;
  Best Local Similarity   95.7%;  Pred. No. 7.5e-95;
  Matches  379;  Conservative    0;  Mismatches    5;  Indels   12;  Gaps    1;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405


RESULT 30
US-09-895-793-822
; Sequence 822, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-793-822

  Query Match             89.4%;  Score 354;  DB 9;  Length 675;
  Best Local Similarity   95.7%;  Pred. No. 7.5e-95;
  Matches  379;  Conservative    0;  Mismatches    5;  Indels   12;  Gaps    1;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405


Search completed: April 30, 2004, 07:45:33
Job time : 295 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 30, 2004, 05:13:09 ; Search time 64 Seconds
(without alignments)
3433.759 Million cell updates/sec

Title: US-09-684-215B-3
Perfect score: 396
Sequence: 1 acggccgcgtccgataactt.....tggccgagggacccccggcc 396

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	396	100.0	702	4	US-09-287-849-27	Sequence 27, Appl
2	396	100.0	2287	4	US-09-223-040-1	Sequence 1, Appli
3	396	100.0	2287	4	US-09-287-849-1	Sequence 1, Appli
4	396	100.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
5	396	100.0	4411529	3	US-09-103-840A-1	Sequence 1, Appli
6	394.4	99.6	447	3	US-08-818-112-4	Sequence 4, Appli
7	394.4	99.6	447	4	US-08-818-111-4	Sequence 4, Appli
8	394.4	99.6	447	4	US-09-056-556-4	Sequence 4, Appli
9	394.4	99.6	447	4	US-09-072-596-4	Sequence 4, Appli
10	394.4	99.6	447	4	US-09-072-967-4	Sequence 4, Appli
11	391.2	98.8	1872	3	US-08-818-112-17	Sequence 17, Appl
12	391.2	98.8	1872	4	US-08-818-111-17	Sequence 17, Appl
13	391.2	98.8	1872	4	US-09-056-556-17	Sequence 17, Appl
14	391.2	98.8	1872	4	US-09-072-596-17	Sequence 17, Appl
15	391.2	98.8	1872	4	US-09-072-967-17	Sequence 17, Appl
16	354	89.4	675	4	US-09-636-215-822	Sequence 822, App
17	354	89.4	675	4	US-09-685-166A-822	Sequence 822, App
18	354	89.4	822	4	US-09-736-457-1862	Sequence 1862, Ap
19	354	89.4	900	4	US-09-643-597-353	Sequence 353, App
20	354	89.4	900	4	US-09-606-421B-353	Sequence 353, App
21	354	89.4	915	4	US-09-636-215-834	Sequence 834, App
22	354	89.4	915	4	US-09-685-166A-834	Sequence 834, App
23	354	89.4	945	4	US-09-736-457-1861	Sequence 1861, Ap
24	354	89.4	1012	4	US-09-643-597-351	Sequence 351, App
25	354	89.4	1012	4	US-09-606-421B-351	Sequence 351, App
26	354	89.4	1203	4	US-09-636-215-851	Sequence 851, App
27	354	89.4	1203	4	US-09-685-166A-851	Sequence 851, App
28	354	89.4	1464	4	US-09-620-412C-348	Sequence 348, App
29	354	89.4	1464	4	US-09-598-419-348	Sequence 348, App
30	354	89.4	1557	4	US-09-620-412C-332	Sequence 332, App
31	354	89.4	1557	4	US-09-598-419-332	Sequence 332, App
32	354	89.4	1578	4	US-09-556-877-188	Sequence 188, App
33	354	89.4	1578	4	US-09-620-412C-188	Sequence 188, App
34	354	89.4	1578	4	US-09-598-419-188	Sequence 188, App
35	354	89.4	1752	4	US-09-620-412C-352	Sequence 352, App
36	354	89.4	1752	4	US-09-598-419-352	Sequence 352, App
37	354	89.4	1758	4	US-09-620-412C-336	Sequence 336, App
38	354	89.4	1758	4	US-09-598-419-336	Sequence 336, App
39	354	89.4	1860	4	US-09-620-412C-308	Sequence 308, App
40	354	89.4	1860	4	US-09-598-419-308	Sequence 308, App
41	354	89.4	1896	4	US-09-620-412C-324	Sequence 324, App
42	354	89.4	1896	4	US-09-598-419-324	Sequence 324, App
43	354	89.4	1941	4	US-09-620-412C-316	Sequence 316, App
44	354	89.4	1941	4	US-09-598-419-316	Sequence 316, App
45	354	89.4	1965	4	US-09-620-412C-340	Sequence 340, App

ALIGNMENTS

RESULT 1
US-09-287-849-27
; Sequence 27, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1
; NAME/KEY: CDS
; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-09-287-849-27

Query Match 100.0%; Score 396; DB 4; Length 702;
Best Local Similarity 100.0%; Pred. No. 3.2e-95;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 2
US-09-223-040-1
; Sequence 1, Application US/09223040
; Patent No. 6544522
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009010US
; CURRENT APPLICATION NUMBER: US/09/223,040
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
; OTHER INFORMATION: protein Ra12-TbH9-Ra35
; NAME/KEY: modified_base
; LOCATION: (30)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: modified_base
; LOCATION: (33)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
; NAME/KEY: modified_base
; LOCATION: (2270)
; OTHER INFORMATION: n = g, a, c or t
US-09-223-040-1

Query Match 100.0%; Score 396; DB 4; Length 2287;
Best Local Similarity 100.0%; Pred. No. 4.1e-95;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

RESULT 3
US-09-287-849-1
; Sequence 1, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
; OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39
; OTHER INFORMATION: fusion)
; NAME/KEY: modified_base
; LOCATION: (30)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: modified_base
; LOCATION: (33)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
; NAME/KEY: modified_base
; LOCATION: (2270)

```
;   OTHER INFORMATION: n = g, a, c or t
US-09-287-849-1

  Query Match             100.0%;  Score 396;  DB 4;  Length 2287;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-95;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458


RESULT 4
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

  Query Match             100.0%;  Score 396;  DB 3;  Length 4403765;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-94;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151984 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 152043

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152044 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 152103

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152104 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 152163

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152164 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 152223

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152224 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 152283

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152284 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 152343

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db     152344 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 152379


RESULT 5
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

  Query Match             100.0%;  Score 396;  DB 3;  Length 4411529;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-94;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151815 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 151874

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151875 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 151934

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151935 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 151994

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151995 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 152054

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152055 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 152114

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152115 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 152174
```

```
Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db     152175 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 152210
```

RESULT 6
US-08-818-112-4
; Sequence 4, Application US/08818112
; Patent No. 6290969
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,112
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-818-112-4

Query Match 99.6%; Score 394.4; DB 3; Length 447;
Best Local Similarity 99.7%; Pred. No. 7.6e-95;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

RESULT 7
US-08-818-111-4
; Sequence 4, Application US/08818111
; Patent No. 6338852
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,111
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-818-111-4

Query Match 99.6%; Score 394.4; DB 4; Length 447;
Best Local Similarity 99.7%; Pred. No. 7.6e-95;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
```

```
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406


RESULT 8
US-09-056-556-4
; Sequence 4, Application US/09056556
; Patent No. 6350456
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND           TREATM
;    NUMBER OF SEQUENCES:  241
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,556
;      FILING DATE:  07-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.457
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-056-556-4

  Query Match             99.6%;  Score 394.4;  DB 4;  Length 447;
  Best Local Similarity   99.7%;  Pred. No. 7.6e-95;
  Matches  395;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190
```

```
Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406


RESULT 9
US-09-072-596-4
; Sequence 4, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-4

  Query Match             99.6%;  Score 394.4;  DB 4;  Length 447;
  Best Local Similarity   99.7%;  Pred. No. 7.6e-95;
  Matches  395;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
                |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
                |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

```
RESULT 10
US-09-072-967-4
; Sequence 4, Application US/09072967
; Patent No. 6592877
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  355
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,967
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-967-4

  Query Match             99.6%;  Score 394.4;  DB 4;  Length 447;
```

```
  Best Local Similarity   99.7%;  Pred. No. 7.6e-95;
  Matches  395;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
                |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
                |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

```
RESULT 11
US-08-818-112-17
; Sequence 17, Application US/08818112
; Patent No. 6290969
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  153
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,112
;      FILING DATE:  13-MAR-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-818-112-17

  Query Match             98.8%;  Score 391.2;  DB 3;  Length 1872;
  Best Local Similarity   99.2%;  Pred. No. 7.2e-94;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153


RESULT 12
US-08-818-111-17
; Sequence 17, Application US/08818111
; Patent No. 6338852
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  148
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,111
;      FILING DATE:  13-MAR-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-818-111-17

  Query Match             98.8%;  Score 391.2;  DB 4;  Length 1872;
  Best Local Similarity   99.2%;  Pred. No. 7.2e-94;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153


RESULT 13
US-09-056-556-17
; Sequence 17, Application US/09056556
; Patent No. 6350456
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND          TREAT
;    NUMBER OF SEQUENCES:  241
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,556
;      FILING DATE:  07-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.457
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-056-556-17

  Query Match             98.8%;  Score 391.2;  DB 4;  Length 1872;
  Best Local Similarity   99.2%;  Pred. No. 7.2e-94;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153


RESULT 14
US-09-072-596-17
; Sequence 17, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-17

  Query Match             98.8%;  Score 391.2;  DB 4;  Length 1872;
  Best Local Similarity   99.2%;  Pred. No. 7.2e-94;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153


RESULT 15
US-09-072-967-17
; Sequence 17, Application US/09072967
; Patent No. 6592877
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  355
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
```

```
;       COUNTRY:  USA
;       ZIP:  98104-7092
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/072,967
;       FILING DATE:  05-MAY-1998
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Maki, David J.
;       REGISTRATION NUMBER:  31,392
;       REFERENCE/DOCKET NUMBER:  210121.411C9
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  17:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1872 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-09-072-967-17

  Query Match             98.8%;  Score 391.2;  DB 4;  Length 1872;
  Best Local Similarity   99.2%;  Pred. No. 7.2e-94;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153


RESULT 16
US-09-636-215-822
; Sequence 822, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.42717C17
;  CURRENT APPLICATION NUMBER: US/09/636,215
;  CURRENT FILING DATE:  2000-08-10
;  NUMBER OF SEQ ID NOS:  852
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-636-215-822

  Query Match             89.4%;  Score 354;  DB 4;  Length 675;
  Best Local Similarity   95.7%;  Pred. No. 3.4e-84;
  Matches  379;  Conservative    0;  Mismatches    5;  Indels   12;  Gaps    1;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405


RESULT 17
US-09-685-166A-822
; Sequence 822, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
```

; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-822

Query Match 89.4%; Score 354; DB 4; Length 675;
Best Local Similarity 95.7%; Pred. No. 3.4e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||              |   ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 18
US-09-736-457-1862
; Sequence 1862, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-1862

Query Match 89.4%; Score 354; DB 4; Length 822;
Best Local Similarity 95.7%; Pred. No. 3.5e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||              |   ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 19
US-09-643-597-353
; Sequence 353, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-643-597-353

Query Match 89.4%; Score 354; DB 4; Length 900;
Best Local Similarity 95.7%; Pred. No. 3.6e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81
```

```
Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||              |   ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 20
US-09-606-421B-353
; Sequence 353, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-421B-353

Query Match 89.4%; Score 354; DB 4; Length 900;
Best Local Similarity 95.7%; Pred. No. 3.6e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||              |   ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 21
US-09-636-215-834
; Sequence 834, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-636-215-834

Query Match 89.4%; Score 354; DB 4; Length 915;
Best Local Similarity 95.7%; Pred. No. 3.6e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||              |   ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
                ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405


RESULT 22
US-09-685-166A-834
; Sequence 834, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C21
;  CURRENT APPLICATION NUMBER: US/09/685,166A
;  CURRENT FILING DATE:  2000-10-10
;  NUMBER OF SEQ ID NOS: 898
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
;   LENGTH: 915
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-685-166A-834

  Query Match             89.4%;  Score 354;  DB 4;  Length 915;
  Best Local Similarity   95.7%;  Pred. No. 3.6e-84;
  Matches  379;  Conservative    0;  Mismatches    5;  Indels   12;  Gaps    1;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||  |           |  ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405


RESULT 23
US-09-736-457-1861
; Sequence 1861, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1861
;   LENGTH: 945
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-736-457-1861

  Query Match             89.4%;  Score 354;  DB 4;  Length 945;
  Best Local Similarity   95.7%;  Pred. No. 3.7e-84;
  Matches  379;  Conservative    0;  Mismatches    5;  Indels   12;  Gaps    1;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||  |           |  ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405


RESULT 24
US-09-643-597-351
; Sequence 351, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
```

```
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
;  CURRENT APPLICATION NUMBER: US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 351
;   LENGTH: 1012
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-643-597-351

  Query Match             89.4%;  Score 354;  DB 4;  Length 1012;
  Best Local Similarity   95.7%;  Pred. No. 3.7e-84;
  Matches  379;  Conservative    0;  Mismatches    5;  Indels   12;  Gaps    1;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||           |   ||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATC-----------AAGCTTCCCACCGTTCAT 174

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 234

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 294

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 354

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 414

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        415 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 450


RESULT 25
US-09-606-421B-351
; Sequence 351, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C9
;  CURRENT APPLICATION NUMBER: US/09/606,421B
;  CURRENT FILING DATE:  2000-06-28
;  NUMBER OF SEQ ID NOS: 358
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 351
;   LENGTH: 1012
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-606-421B-351

  Query Match             89.4%;  Score 354;  DB 4;  Length 1012;
  Best Local Similarity   95.7%;  Pred. No. 3.7e-84;
  Matches  379;  Conservative    0;  Mismatches    5;  Indels   12;  Gaps    1;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||           |   ||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATC-----------AAGCTTCCCACCGTTCAT 174

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 234

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 294

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 354

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 414

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        415 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 450


RESULT 26
US-09-636-215-851
; Sequence 851, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.42717C17
;  CURRENT APPLICATION NUMBER: US/09/636,215
;  CURRENT FILING DATE:  2000-08-10
;  NUMBER OF SEQ ID NOS:  852
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 851
;   LENGTH: 1203
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

US-09-636-215-851

Query Match 89.4%; Score 354; DB 4; Length 1203;
Best Local Similarity 95.7%; Pred. No. 3.8e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 27
US-09-685-166A-851
; Sequence 851, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-851

Query Match 89.4%; Score 354; DB 4; Length 1203;
Best Local Similarity 95.7%; Pred. No. 3.8e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 28
US-09-620-412C-348
; Sequence 348, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 348
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-620-412C-348

Query Match 89.4%; Score 354; DB 4; Length 1464;
Best Local Similarity 95.7%; Pred. No. 4e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC------------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309
```

```
Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405


RESULT 29
US-09-598-419-348
; Sequence 348, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 348
;   LENGTH: 1464
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-348

  Query Match             89.4%;  Score 354;  DB 4;  Length 1464;
  Best Local Similarity   95.7%;  Pred. No. 4e-84;
  Matches  379;  Conservative    0;  Mismatches    5;  Indels   12;  Gaps    1;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||           |   ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC-----------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405


RESULT 30
US-09-620-412C-332
; Sequence 332, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 332
;   LENGTH: 1557
;   TYPE: DNA
;   ORGANISM: Chlymadia trachomatis
US-09-620-412C-332

  Query Match             89.4%;  Score 354;  DB 4;  Length 1557;
  Best Local Similarity   95.7%;  Pred. No. 4.1e-84;
  Matches  379;  Conservative    0;  Mismatches    5;  Indels   12;  Gaps    1;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||||           |   ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC-----------AAGCTTCCCACCGTTCAT 129

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405


Search completed: April 30, 2004, 07:46:54
Job time : 72 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 22:25:27 ; Search time 2585.37 Seconds
(without alignments)
1524.660 Million cell updates/sec

Title: US-09-684-215B-4
Perfect score: 675
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....QTKSGGTRTGNVTLAEGPPA 132

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061305_13200/app_query.fasta_1.1180
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_5167_@runat_29042004_061305_13200 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		145.5	21.6	289	29	U82114	U82114 U82114 orde
2	c	126.5	18.7	1438	28	BH770798	BH770798 LLMGtag54
3	c	122	18.1	959	28	BZ549048	BZ549048 pacs1-60_
4	c	120.5	17.9	726	14	CB679186	CB679186 OSJNEf02F
5		117.5	17.4	590	14	CB925602	CB925602 ABA1_22_F
6	c	115.5	17.1	807	14	CA228900	CA228900 SCQGFL305
7		114	16.9	603	28	AZ934428	AZ934428 BJ__Ba000
8	c	113	16.7	758	13	BQ514888	BQ514888 EST622303
9		108	16.0	982	14	CD080132	CD080132 MA3-9999U
10	c	106.5	15.8	1033	28	BZ561390	BZ561390 pacs2-164
11		103.5	15.3	719	28	AZ933900	AZ933900 BJ__Ba000
12	c	103.5	15.3	947	28	BZ549047	BZ549047 pacs1-60_
13		103	15.3	726	28	AQ989479	AQ989479 Rfc00025
14		102	15.1	765	9	AJ558965	AJ558965 AJ558965
15	c	100.5	14.9	859	14	CA480614	CA480614 AGENCOURT
16		99	14.7	423	9	AI597611	AI597611 tn15f02.x
17	c	97.5	14.4	499	12	BI350520	BI350520 fr32b05.y
18	c	97.5	14.4	870	14	CA473840	CA473840 AGENCOURT
19		97	14.4	444	13	CA148171	CA148171 SCEZRZ101
20		97	14.4	543	12	BI721127	BI721127 1031054B1
21		96	14.2	546	10	AW285510	AW285510 LG1_241_E
22		96	14.2	551	10	AW285527	AW285527 LG1_241_G
23	c	96	14.2	1093	13	BU557763	BU557763 AGENCOURT
24		96	14.2	1394	28	BZ576076	BZ576076 msh2_4761
25	c	95.5	14.1	600	14	CF787130	CF787130 854999 MA
26	c	95.5	14.1	647	13	BX674871	BX674871 BX674871
27		95.5	14.1	677	12	BI378928	BI378928 BFLG1_000
28		95	14.1	881	14	CK161112	CK161112 FGAS04279
29	c	95	14.1	909	14	CA975497	CA975497 AGENCOURT
30	c	94.5	14.0	728	12	BJ285991	BJ285991 BJ285991
31	c	94.5	14.0	1029	12	BI457360	BI457360 603185681
32		94	13.9	434	28	BH638292	BH638292 1008021F0
33		94	13.9	672	14	CF844326	CF844326 psHB027xJ
34		94	13.9	1321	13	BU509638	BU509638 AGENCOURT
35		93.5	13.9	600	28	CC345457	CC345457 OGQAF70TH
36		93.5	13.9	766	14	CF997044	CF997044 AGENCOURT
37	c	93.5	13.9	907	14	CA471201	CA471201 AGENCOURT
38	c	93.5	13.9	913	14	CF265501	CF265501 AGENCOURT
39	c	93	13.8	689	12	BM448307	BM448307 DSA024D06
40	c	93	13.8	951	14	CA474469	CA474469 AGENCOURT
41		92.5	13.7	452	12	BM874057	BM874057 laa07a02.
42	c	92.5	13.7	709	29	CG895905	CG895905 pastbac00
43	c	92.5	13.7	860	14	CA471291	CA471291 AGENCOURT
44	c	92.5	13.7	949	14	CF243195	CF243195 AGENCOURT
45		92.5	13.7	1117	14	CK208964	CK208964 FGAS02069

ALIGNMENTS

RESULT 1
U82114
LOCUS U82114 289 bp DNA linear GSS 16-FEB-2001
DEFINITION U82114 ordered cosmid library Mycobacterium leprae genomic clone cosmid L-373; contig 64, genomic survey sequence.
ACCESSION U82114
VERSION U82114.1 GI:3647212
KEYWORDS GSS.
SOURCE Mycobacterium leprae
ORGANISM Mycobacterium leprae
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
REFERENCE 1 (bases 1 to 289)

```
  AUTHORS   Silbaq,F.S., Cho,S.N., Cole,S.T. and Brennan,P.J.
  TITLE     Characterization of a 34-kilodalton protein of Mycobacterium leprae
            that is isologous to the immunodominant 34-kilodalton antigen of
            Mycobacterium paratuberculosis
  JOURNAL   Infect. Immun. 66 (11), 5576-5579 (1998)
  MEDLINE   99003183
   PUBMED   9784577
COMMENT     Contact: Silbaq FS
            Microbiology
            Colorado State University
            Fort Collins, CO 80523, USA
            Eiglmeier,K., Honore,N., Woods,S.A., Caudron,B. and Cole,S.T. Use
            of an ordered cosmid library to deduce the genomic organization of
            Mycobacterium leprae. Mol. Microbiol. 7 (2), 197-206 (1993)
            Class: unknown.
FEATURES             Location/Qualifiers
     source          1. .289
                     /organism="Mycobacterium leprae"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1769"
                     /clone="cosmid L-373; contig 64"
                     /clone_lib="ordered cosmid library"
ORIGIN

Alignment Scores:
Pred. No.:                 8.75e-05        Length:          289
Score:                     145.50          Matches:         35
Percent Similarity:        58.82%          Conservative:    15
Best Local Similarity:     41.18%          Mismatches:      34
Query Match:               21.56%          Indels:          1
DB:                        29              Gaps:            1

US-09-684-215B-4 (1-132) x U82114 (1-289)

Qy          45 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsn---GlyAlaArgValGlnArg 63
               |||   ||||||:::   |||      :::   |||      ||||||:::|||
Db           7 GCGTCGCTGGGTGTACAGGTGGCCACCGACAAAGGCACCCCGGGCGCCAAAGTTATGGAC 66

Qy          64 ValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAla 83
               ||||||           ||||||:::      :::      |||   |||:::|||
Db          67 GTCGTGGCCGGTGGTGCGGCCGCGAATGCCGCCGTTCCCAAGGGGGTCGTTCTCACTAAG 126

Qy          84 ValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHis 103
               ||||||          |||:::||||||   |||:::      |||:::
Db         127 GTCGACGACCGCCTGATCAGTAGCGCCGACGCGTTGGTCGCTGCCGTGCGGTCCAAGGCA 186

Qy         104 ProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsn 123
               |||||||||   :::|||:::|||:::|||   :::||||||   :::||||||
Db         187 CCCGGTGACAAAGTGTCGCTGACCTATCAGGATCAGTCTGGTAGCAGTCGCACGGTTCAG 246

Qy         124 ValThrLeuAlaGlu 128
               |||||||||   :::
Db         247 GTCACACTCGGCAAG 261

RESULT 2
BH770798/c
LOCUS       BH770798                1438 bp    DNA     linear   GSS 01-MAY-2002
DEFINITION  LLMGtag541 MG1363 Random Sequence Tag Library Lactococcus lactis
            subsp. cremoris genomic, genomic survey sequence.
ACCESSION   BH770798
VERSION     BH770798.1  GI:20373755
KEYWORDS    GSS.
SOURCE      Lactococcus lactis subsp. cremoris
  ORGANISM  Lactococcus lactis subsp. cremoris
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Lactococcus.
REFERENCE   1  (bases 1 to 1438)
  AUTHORS   Bolotin,A., Ehrlich,S.D. and Sorokin,A.
  TITLE     Studies of genomes of dairy bacteria Lactococcus lactis
  JOURNAL   Sci. Aliments (2002) In press
COMMENT     Contact: Sorokin A
            Genetique Microbienne
            INRA
            CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
            Tel: 33 1 34 65 25 16
            Fax: 33 1 34 65 25 21
            Email: sorokine@jouy.inra.fr
            best homologue in strain IL1403 is htrA (95%)
            Class: shotgun
            High quality sequence start: 30
            High quality sequence stop: 1408.
FEATURES             Location/Qualifiers
     source          1. .1438
                     /organism="Lactococcus lactis subsp. cremoris"
                     /mol_type="genomic DNA"
                     /strain="MG1363"
                     /sub_species="cremoris"
                     /db_xref="taxon:1359"
                     /clone_lib="MG1363 Random Sequence Tag Library"
                     /note="Vector: pSGMU2; Site_1: SmaI; Library of
                     chromosomal fragments of L.lactis strain MG1363 was
                     prepared by partial AluI digestion or by sonication."
ORIGIN

Alignment Scores:
Pred. No.:                 0.0656          Length:          1438
Score:                     126.50          Matches:         40
Percent Similarity:        45.45%          Conservative:    20
Best Local Similarity:     30.30%          Mismatches:      49
Query Match:               18.74%          Indels:          23
DB:                        28              Gaps:            5

US-09-684-215B-4 (1-132) x BH770798 (1-1438)

Qy          13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArgSer 32
               |||   ||||||||||||||          :::   |||     ::::::   :::
Db         479 GGTTTAGGTTTTGCTATCCCATCTAATGATGTGGTAAACATCATTAATAAACTTGAAACT 420

Qy          33 GlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValVal 52
               |||                 |||       |||   ||||||:::   :::|||
Db         419 GATGGT--------------AAGATTTCACGTCCTGCC---TTAGGTATTCGTATGGTT 378

Qy          53 Asp-----------------AsnAsnGlyAsn-------------------------- 57
               |||                 |||:::   :::
Db         377 GACCTATCTCAATTATCAACAAATGATAGTTCTCAACTGAAATTACCTAGCAGCGTAACT 318

Qy          58 ---GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSer 76
                  |||       |||       |||                 ||||||:::   |||:::
Db         317 GGTGGAGTGGTTGTCTACTCTGTTCAAGCGGGTCTTTCTGCTGCCACAGCTGGTCTGAAA 258

Qy          77 ThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAla 96
               |||||||||||||||   |||           :::   |||:::|||   :::
Db         257 GCTGGCGATGTGATAACGAAGGTGGGAGATACCGCCGTTACTTCATCAACAGACTTACAA 198

Qy          97 AspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSer 116
               ||||||      |||:::      |||   :::   ||||||:::      :::
Db         197 AGTGCTCTTTACTCACACAATATTAATGATACTGTGAAAGTCACCTACTAC---CGTGAT 141

Qy         117 GlyGlyThrArgThrGlyAsnValThrLeuAlaGlu 128
               |||   :::   |||   ||||||   |||::::::
Db         140 GGTAAATCAGCCACAGCAAATGTCAAACTCTCTAAA 105

RESULT 3
BZ549048/c
LOCUS       BZ549048                  959 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  pacs1-60_1644.s2 pacs1-60 Pseudomonas aeruginosa genomic clone
            pacs1-60_1644, genomic survey sequence.
ACCESSION   BZ549048
VERSION     BZ549048.1  GI:27152629
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 959)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
```

```
          Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE   Whole-Genome-Sequence variation among multiple isolates of
          Psedomonas aeruginosa library
  JOURNAL J. Bacteriol. (2002) In press
COMMENT   Contact: Chris K. Raymond
          Genome Center
          University of Washington
          Box 352145, Seattle, WA 98105-2145, USA
          Tel: 2062216954
          Fax: 2066857244
          Email: craymond@u.washington.edu
          Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .959
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="1-60"
                     /db_xref="taxon:287"
                     /clone="pacs1-60_1644"
                     /clone_lib="pacs1-60"
                     /note="clinical isolate 1-60 Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:                0.101          Length:          959
Score:                    122.00         Matches:         40
Percent Similarity:       47.29%         Conservative:    21
Best Local Similarity:    31.01%         Mismatches:      44
Query Match:              18.07%         Indels:          24
DB:                       28             Gaps:            5

US-09-684-215B-4 (1-132) x BZ549048 (1-959)

Qy          4 SerAspAsnPheGlnLeuSerGlnGly------GlyGlnGlyPheAlaIleProIleGly 21
              |||       |||       |||  |||       |||      ||||||||||||||||
Db        378 TCGCAGATCTTCACCCGTTCCGGCGGCTTCATGGGCCTGTCCTTCGCCATTCCGATCGAT 319

Qy         22 GlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHisIle 41
                ||:::   :::|||   |||::::::      |||           |||
Db        318 GTCGCGCTGAACGTCGCCGACCAGTTGAAGAAAGCCGGC---------AAGGTCAGTCGC 268

Qy         42 GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsn------------------ 55
              |||         :::||||||:::   :::   :::   |||
Db        267 GGC---------TGGCTGGGCGTGGTGATCCAGGAAGTGAACAAGGATCTCGCCGAGTCC 217

Qy         56 ---------------GlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAla 70
                          :::|||||||   |||   ::::::|||         ||||||
Db        216 TACGGCCTCGACAAGCCGTCCGGCGCGCTGGTGGCGCAACTGGTGGAAGACGGCCCGGCG 157

Qy         71 AlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsn 90
              |||      |||:::     |||||||||||   :::::::::|||         ||||||
Db        156 GCCAAGGGTGGCCTGCAGGTGGGCGATGTGATCATAAGCTTGAACGGCCAGTCGATCAAC 97

Qy         91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110
                 :::      :::         :::              |||||||||   |||::::::
Db         96 GAGTCCGCCGACCTGCCGCACCTGGTGGGCAACATGAAGCCGGGCGACAAGATCAACCTG 37

Qy        111 ThrTrpGlnThrLysSerGlyGlyThr 119
                           ||||||:::
Db         36 ---------------GACGGGGGATCC 25

RESULT 4
CB679186/c
LOCUS       CB679186                 726 bp    mRNA    linear   EST 09-APR-2003
DEFINITION  OSJNEf02F19.r OSJNEf Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEf02F19 3', mRNA sequence.
ACCESSION   CB679186
VERSION     CB679186.1  GI:29682911
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 02  row: F  column: 19
            Seq primer: gga aac agc tat gac cat g.
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEf02F19"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEf"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Uninfected Control"
ORIGIN

Alignment Scores:
Pred. No.:                0.0952         Length:          726
Score:                    120.50         Matches:         41
Percent Similarity:       44.03%         Conservative:    18
Best Local Similarity:    30.60%         Mismatches:      50
Query Match:              17.85%         Indels:          25
DB:                       14             Gaps:            4

US-09-684-215B-4 (1-132) x CB679186 (1-726)

Qy         10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla------ 27
              :::      |||   ||||||||||||||||           :::   ||||||
Db        579 ACATCTGCTGGTGTTGGTTTTGCCATCCCGTCATCAACTGTTCTGAAAATAGCTCCTCAG 520

Qy         28 ------------GlyGlnIleArgSerGlyGlyGlySerProThrValHisIleGlyPro 43
                         |||::::::|||      |||   :::                  |||
Db        519 TTAATTCAGTTTGGAAAAGTTCGCCGTGCTGGCTTGAATGTGGAGTTTGCTCCAGATCCA 460

Qy         44 ThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArg 63
              |||:::                  |||          |||:::   :::   :::
Db        459 ATCGCATAT---------------CAGCTTAATGTTCGCACTGGATCTATAATATTGCAG 415

Qy         64 ValValGlySerAlaProAlaAlaSerLeuGly--------------------------- 74
              |||   |||   :::   ||||||    |||
Db        414 GTTCCTGGGGGCAGTGCTGCAGCCAAAGCTGGTCTTGTTCCTACCAGTAGGGGTTTTGCT 355

Qy         75 ------IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92
                    |||      ||||||||||||   ||||||||||   ||||||
Db        354 GGTACCATTGTTCTTGGTGATGTCATTGTTGCCGTGGACGGTAAACCTATCAAAGGCAAA 295

Qy         93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 112
              :::   ::::::      |||:::   :::      ||||||   :::|||:::|||
Db        294 TCTGATCTATCGAGGGTTCTGGATGCTTATGGCGTTGGAGACAAGGTGAGCTTGACAATC 235

Qy        113 GlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 126
              |||         |||      :::      |||||||||
Db        234 CAAAGA---------GGCGCTGAATCTCTGGAGGTAACCTTG 202
```

```
RESULT 5
CB925602
LOCUS       CB925602                590 bp    mRNA    linear   EST 28-APR-2003
DEFINITION  ABA1_22_F06.b1_A012 Abscisic acid-treated seedlings Sorghum bicolor
            cDNA clone ABA1_22_F06_A012 3', mRNA sequence.
ACCESSION   CB925602
VERSION     CB925602.1  GI:30161873
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Cordonnier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S.,
            Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Buchanan,C.D.,
            Eastman,A. and Pratt,L.H.
  TITLE     An EST database from Sorghum: ABA1-treated seedlings
  JOURNAL   Unpublished (2003)
COMMENT     Other_ESTs: ABA1_22_F06.g1_A012
            Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
            the Human Genome Center, University of Tokyo Institute of Medical
            Science; plant material and RNA prepared at Texas A & M University;
            sequencing done in the Laboratory for Genomics and Bioinformatics,
            University of Georgia. Sequence ends have been trimmed to exclude
            vector and regions below Phred quality 16. Three-prime sequences
            are presented as their reverse complement and have been trimmed to
            exclude polyA.
            Seq primer: Sug3 (CGACCTGCAGCTCGAGCACA)
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /cultivar="IS3620C"
                     /db_xref="taxon:4558"
                     /clone="ABA1_22_F06_A012"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="Abscisic acid-treated seedlings"
                     /note="Vector: pME18S-FL3; Site_1: XhoI; Site_2: XhoI; The
                     library was prepared from polyA+ RNA from seedlings grown
                     in hydroponic culture. After 12 days, medium was
                     supplemented with 1 mM abscisic acid (ABA), while leaves
                     were misted with a solution of 1 mM ABA. Roots and leaves
                     were harvested after 3, 6, 12, and 24 hr and material from
                     all time points was combined prior to RNA isolation.
                     Double-stranded cDNA was cloned unidirectionally into
                     different DraIII sites of the pME18S-FL3 vector (5-prime
                     DraIII site is CACTGTGTG, 3-prime DraIII site is
                     CACCATGTG). XhoI excises the cDNA insert."
ORIGIN

Alignment Scores:
Pred. No.:                0.139          Length:         590
Score:                    117.50         Matches:        37
Percent Similarity:       44.19%         Conservative:   20
Best Local Similarity:    28.68%         Mismatches:     43
Query Match:              17.41%         Indels:         29
DB:                       14             Gaps:           4

US-09-684-215B-4 (1-132) x CB925602 (1-590)

Qy          15 GlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla----------------Gly 28
               ||||||||||||||||           :::   ||||||                 |||
Db           6 GGCTTTGCTATTCCATCATCAACTGTACTTAAAATCGCTCCTCAGTTAATTCAGTTCGGA 65

Qy          29 GlnIleArgSerGlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGly 48
               ::::::|||      |||   :::                     |||   |||:::
Db          66 AAAGTTCGTCGTGCTGGCTTGAATGTGGACTTCGCTCCGGATCCAATTGCATAT------ 119

Qy          49 LeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAla 68
                        |||       |||||||||  :::   :::|||   |||   :::
Db         120 ---------CAGCTTAATGTTCGCAACGGAGCTCTTATACTTAAGGTACCTGGGGGCAGT 170

Qy          69 ProAlaAlaSerLeuGlyIleSer--------------------------------Thr 77
               ||||||      |||::::::
Db         171 GCTGCAGCCAAAGCAGGTCTTGCTCCAACCGGCAGGGGTTTTGCTGGTAATATTGTTCTG 230

Qy          78 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 97
               ||||||:::|||   ||||||||||||   |||:::         :::   :::
Db         231 GGTGATATCATCGTTGCAGTGGACGGCAAACCTGTTAAGGGCAAATCTGACCTGCTGAGG 290

Qy          98 AlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGly 117
                  |||:::   :::      ||||||   :::::::::|||
Db         291 GTTCTGGATGACTATGGCGTCGGAGATCAGGTGACCTTGACA----------------- 332

Qy         118 GlyThrArgThrGlyAsnValThrLeu 126
                  |||   |||:::   ||||||
Db         333 ---ATCCGGCGAGGCTCAGAAACCCTT 356

RESULT 6
CA228900/c
LOCUS       CA228900                807 bp    mRNA    linear   EST 25-SEP-2003
DEFINITION  SCQGFL3058G06.g Saccharum officinarum FL3 Saccharum officinarum
            cDNA clone SCQGFL3058G06 5', mRNA sequence.
ACCESSION   CA228900
VERSION     CA228900.1  GI:35290878
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE   1  (bases 1 to 807)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 058  row: G  column: 06
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .807
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCQGFL3058G06"
                     /lab_host="DH10B"
                     /clone_lib="Saccharum officinarum FL3"
                     /note="Organ: Base of developing inflorescence (5cm-long);
                     Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [Base of
                     developing  inflorescence (5cm-long)]. cDNA was prepared
                     from polyA+ mRNA using  SuperScript Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were  fractionated
                     in a sepharose CL-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details  of each source of RNA and library
                     construction can be obtained at
                     http://sucest.lad.ic.unicamp.br/public"
```

```
ORIGIN

Alignment Scores:
Pred. No.:                0.344          Length:          807
Score:                    115.50         Matches:         38
Percent Similarity:       43.28%         Conservative:    20
Best Local Similarity:    28.36%         Mismatches:      47
Query Match:              17.11%         Indels:          29
DB:                       14             Gaps:            4

US-09-684-215B-4 (1-132) x CA228900 (1-807)

Qy          10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla------ 27
               :::     |||  ||||||||||||          :::   ||||||
Db         569 ACATCTGCTGGTGTTGGCTTCGCTATCTCATCATCAACTGTGCTTAAAATCGCTCCTCAG 510

Qy          28 ------------GlyGlnIleArgSerGlyGlyGlySerProThrValHisIleGlyPro 43
                        |||:::::::|||      |||   :::                 |||
Db         509 TTAATTCAGTTTGGAAAAGTTCGTCGTGCTGGCTTGAATGTGGACTTTGCTCCAGATCCA 450

Qy          44 ThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArg 63
               |||:::                  |||      |||||||||   :::   :::
Db         449 ATTGCGTAT---------------CAGCTTAATGTTCGCAACGGAGCTCTTATACTTAAG 405

Qy          64 ValValGlySerAlaProAlaAlaSerLeuGlyIleSer--------------------- 76
               |||   |||   :::   ||||||      |||::::::
Db         404 GTACCTGGGGGCAGTGCTGCAGCCAAGGCAGGCCTTGCTCCGACCGGCAGGGGTTTCGCT 345

Qy          77 ------------ThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92
                        ||||||||||||   ||||||||||   |||:::
Db         344 GGTAACATTGTTCTTGGTGATGTCATCGTTGCAGTGGATGGCAAACCTGTTAAGGGCAAA 285

Qy          93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 112
               :::   :::       |||:::   :::       ||||||  :::::::::|||
Db         284 TCTGACCTGCTGAGGGTTCTGGATGACTATGGCGTCGGGGATCAGGTGACCTTGACA--- 228

Qy         113 GlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 126
                        |||   |||:::   ||||||
Db         227 ------------------ATCCGGCGAGGCTCTGAAACCCTC 204

RESULT 7
AZ934428
LOCUS       AZ934428                603 bp    DNA     linear   GSS 24-APR-2001
DEFINITION  BJ__Ba0002I08r B. japonicum BAC library Bradyrhizobium japonicum
            genomic, genomic survey sequence.
ACCESSION   AZ934428
VERSION     AZ934428.1  GI:13776488
KEYWORDS    GSS.
SOURCE      Bradyrhizobium japonicum
  ORGANISM  Bradyrhizobium japonicum
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE   1  (bases 1 to 603)
  AUTHORS   Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A.,
            Goicoechea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.
  TITLE     A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
            genome
  JOURNAL   Genome Res. 11 (8), 1434-1440 (2001)
  MEDLINE   21376150
   PUBMED   11483585
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Class: BAC ends
            High quality sequence stop: 553.
FEATURES             Location/Qualifiers
     source          1. .603
                     /organism="Bradyrhizobium japonicum"
                     /mol_type="genomic DNA"
                     /strain="USDA110"
                     /db_xref="taxon:375"
                     /lab_host="E. coli"
                     /clone_lib="B. japonicum BAC library"
                     /note="Vector: pIndigo536; Site_1: HindIII"
ORIGIN

Alignment Scores:
Pred. No.:                0.318          Length:          603
Score:                    114.00         Matches:         36
Percent Similarity:       43.44%         Conservative:    17
Best Local Similarity:    29.51%         Mismatches:      57
Query Match:              16.89%         Indels:          12
DB:                       28             Gaps:            4

US-09-684-215B-4 (1-132) x AZ934428 (1-603)

Qy          13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArgSer 32
               |||      |||:::||||||          :::     |||::::::
Db         245 GGCATCGCGTTCTCGATCCCGGCCAACACCGTGAAGACGGTTGTTGCCCAGCTCAAGGAC 304

Qy          33 GlyGly------------GlySerProThrValHisIleGlyProThrAlaPheLeuGly 48
                  |||            |||         |||   |||   |||
Db         305 AAGGGTTCGGTCAGCCGCGGCTGGATCGGCGTGCAGATTCAGCCGGTGACG--------- 355

Qy          49 LeuGlyValValAspAsnAsnGly------------AsnGlyAlaArgValGlnArgVal 64
                  :::   |||:::   |||               ||||||  |||
Db         356 TCGGATATCGCCGACAGCCTCGGCATGAAGAAGGCCGAAGGGGCGCTGGTGGCGGAGCCG 415

Qy          65 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 84
               :::   |||||||||       ||||||   :::|||||||||||||||:::|||
Db         416 CAGGCGAACGGTCCGGCGGCGAAGGCCGGCATCGAGTCCGGCGACGTGATCACCTCGGTC 475

Qy          85 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 104
               :::|||     :::     |||     :::|||     :::     |||     |||
Db         476 AACGGCGAATCCGTCAAGGACGCCCGCGAGCTCGCCCGCACCATCGGCGGCATGGCGCCC 535

Qy         105 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 124
                  ::::::   :::           |||   |||                 |||:::
Db         536 CGTGCGATCGTGAAGCTTAACGTGCTGCACAAG---GGCCAGGACAATGTCGTGAACCTC 592

Qy         125 ThrLeu 126
               ||||||
Db         593 ACCCTC 598

RESULT 8
BQ514888/c
LOCUS       BQ514888                758 bp    mRNA    linear   EST 07-MAR-2003
DEFINITION  EST622303 Generation of a set of potato cDNA clones for microarray
            analyses mixed potato tissues Solanum tuberosum cDNA clone STMIO81
            3' end, mRNA sequence.
ACCESSION   BQ514888
VERSION     BQ514888.1  GI:21373757
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
            Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
            Karamycheva,S.A.
  TITLE     Generation of a set of potato cDNA clones for microarray analyses
  JOURNAL   Unpublished (2002)
COMMENT     Other_ESTs: EST622302
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/
```

```
          Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .758
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec or Binjte"
                     /db_xref="taxon:4113"
                     /clone="STMIO81"
                     /tissue_type="mixed tissues"
                     /lab_host="SOLR"
                     /clone_lib="Generation of a set of potato cDNA clones for
                     microarray analyses mixed potato tissues"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Combination of untreated and Phytophthora
                     infestans-treated libraries of stolons, leaves, leaflets,
                     axillary buds of stem explants, petioles, germinating
                     eyes, tubers, or roots."
ORIGIN

Alignment Scores:
Pred. No.:                0.554          Length:         758
Score:                    113.00         Matches:        34
Percent Similarity:       42.98%         Conservative:   15
Best Local Similarity:    29.82%         Mismatches:     47
Query Match:              16.74%         Indels:         18
DB:                       13             Gaps:           3

US-09-684-215B-4 (1-132) x BQ514888 (1-758)

Qy         10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
              :::   |||  ||||||||||||||||          :::  :::     |||
Db        592 ACATCAGCAGGTGTTGGATTTGCAATCCCTTCTTCAACTGTGTTGAGAGTTGTGCCCCAG 533

Qy         30 IleArgSerGlyGly------------GlySerProThrValHisIleGlyProThrAla 45
              :::       |||                        :::  |||  |||
Db        532 TTGATCCAATCTGGAAAAGTTCTTCGTGCTGGTTTGAATATTGAAATCGCTCCAGACCTG 473

Qy         46 PheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValVal 65
                  :::            |||      |||||||||  |||      |||
Db        472 ---------ATTGCCAACCAACTTAATGTTCGAAATGGAGCACTGGTTTTGCTGGTACCT 422

Qy         66 GlySerAlaProAlaAlaSerLeuGly------------------------------- 74
              |||::::::|||||||||      |||
Db        421 GGAAATAGTCCTGCAGCGAAAGCGGGACTTCTTCCTACTACCAGGGGTTTTGCAGGAAAT 362

Qy         75 IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAla 94
              |||      ||||||:::|||   |||||||||      |||:::   |||
Db        361 ATAGTGCTTGGCGATATTATTGAAGCAGTGGATGACAAACCTGTTAGGAGTAAAGCAGAG 302

Qy         95 MetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIle 108
              :::      ||||||:::   ::::::   ||||||   :::
Db        301 TTGTATAAAGCCCTGGATAACTATAACATAGGTGATGAAGTT 260

RESULT 9
CD080132
LOCUS       CD080132                982 bp    mRNA    linear   EST 14-SEP-2003
DEFINITION  MA3-9999U-M294-C03-U.B MA3-0001 Schistosoma mansoni cDNA clone
            MA3-9999U-M294-C03.B, mRNA sequence.
ACCESSION   CD080132
VERSION     CD080132.1  GI:34631128
KEYWORDS    EST.
SOURCE      Schistosoma mansoni
  ORGANISM  Schistosoma mansoni
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1  (bases 1 to 982)
  AUTHORS   Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M.,
            Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr.,
            Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F.,
            Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L.,
            Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A.,
            Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A.,
            Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T.,
            Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M.,
            Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
  TITLE     Transcriptome analysis of the acoelomate human parasite Schistosoma
            mansoni
  JOURNAL   Nat. Genet. 35 (2), 148-157 (2003)
  MEDLINE   22879926
COMMENT     Contact: Dr. Sergio Verjovski-Almeida
            Departamento de Bioquimica
            Instituto de Quimica - Universidade de Sao Paulo
            Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP,
            Brasil
            Tel: +55-11-3091-2173
            Fax: +55-11-3091-2186
            Email: verjo@iq.usp.br
            This sequence was derived from the FAPESP Schistosoma mansoni EST
            Genome Project. All sequences in the project were assembled and
            annotated. This entry and all the assembled sequences can be seen
            in the following URL http://bioinfo.iq.usp.br/schisto/
            Plate: MA3-9999U-M294  row: 3  column: C.
FEATURES             Location/Qualifiers
     source          1. .982
                     /organism="Schistosoma mansoni"
                     /mol_type="mRNA"
                     /db_xref="taxon:6183"
                     /clone="MA3-9999U-M294-C03.B"
                     /sex="mixed pool"
                     /dev_stage="adult"
                     /lab_host="Mus musculus"
                     /clone_lib="MA3-0001"
ORIGIN

Alignment Scores:
Pred. No.:                2.5            Length:         982
Score:                    108.00         Matches:        33
Percent Similarity:       37.80%         Conservative:   15
Best Local Similarity:    25.98%         Mismatches:     39
Query Match:              16.00%         Indels:         40
DB:                       14             Gaps:           4

US-09-684-215B-4 (1-132) x CD080132 (1-982)

Qy          6 AsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAla------ 23
              |||      ::::::      |||   |||||||||||||||:::   |||
Db        355 AACGCCATGGTGGCTGGTACAGGCGTGGGATTTGCTATACCTGTTGATCAAGTCCGAAAA 414

Qy         24 ---------MetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
                      :::|||:::             :::       ||||||
Db        415 TTCATACAGTTAGCACTAAAAGCATCCTCAAATACTCGGAGCTCTTCCCCAATATCGTCA 474

Qy         41 IleGlyProThr---------------------------------------------- 44
              ||||||
Db        475 AAATCGCCTACTGACCCATATATTTTAGCAGACAATCCAGCCAATTCGGAAAGTGGAACT 534

Qy         45 ------AlaPheLeuGlyLeu--------------------------------GlyVal 51
                   :::|||||||||                                    :::
Db        535 CAACGTCGATACTTAGGCCTAGTTATGCGGACACTTACTCCAGAACTGGCTTTCGAATTA 594

Qy         52 ValAspAsnAsnGly-----------------AsnGlyAlaArgValGlnArgValVal 65
                        |||                  ||||||      :::      |||:::
Db        595 GCATCTCGCGGTGGACAACATTTTGTAGAGTTGAATGGTGTTCTGATTCATGCTGTATTA 654

Qy         66 GlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAsp 85
              ::::::|||||      |||:::      ||||||||||  |||:::|||
Db        655 AGAAACTCCCCTGCTCANAGAGCGGGATTAAGAGCTGGAGATGTAATTGTTGCCATTGAT 714

Qy         86 GlyAlaProIleAsnSerAla 92
              |||   ||||||   :::|||
Db        715 GGTCTGCCTATAACGAATGCT 735

RESULT 10
BZ561390/c
LOCUS       BZ561390                1033 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  pacs2-164_3239.y3 pacs2-164 Pseudomonas aeruginosa genomic clone
```

```
            pacs2-164_3239, genomic survey sequence.
ACCESSION   BZ561390
VERSION     BZ561390.1  GI:27181349
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 1033)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .1033
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="2-164"
                     /db_xref="taxon:287"
                     /clone="pacs2-164_3239"
                     /clone_lib="pacs2-164"
                     /note="clinical isolate 2-164 Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:              3.78          Length:         1033
Score:                  106.50        Matches:        41
Percent Similarity:     44.20%        Conservative:   20
Best Local Similarity:  29.71%        Mismatches:     56
Query Match:            15.78%        Indels:         21
DB:                     28            Gaps:           4

US-09-684-215B-4 (1-132) x BZ561390 (1-1033)

Qy          4 SerAspAsnPheGlnLeuSerGlnGlyGlyGln------GlyPheAlaIleProIleGly 21
              |||       |||      |||  |||               |||  |||||||||
Db        713 TCCCAGTTTTTCCACCGGTCCGCCGGCTCAATGGCCCTGTCTTTCGGCATTCCGATCGAT 654

Qy         22 GlnAlaMetAlaIle-AlaGlyGlnIleArg--------------------------Se 32
                ||:::   ::: |||  |||::::::                            ::
Db        653 GTCGCGCTGACCGTTCGCCGACCAGTTGAAGAAAGCCCGGCAAGGTCAGTCGCGGCTGGC 594

Qy         32 rGlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeu-GlyLeuGlyValV 52
              :|||   |||   |||          |||            ||| ||||||
Db        593 TGGGCGTGGTGATCCAGGAAGTGAACAAGGATCTCGCCGAGTCCTTCGGCCTC------- 541

Qy         52 alAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaS 72
                 :::      :::|||||   |||   :::::|||         |||||||||
Db        540 -----GACAAGCCGTCCGGCGCGCTGGTGGCGCAACTGGTGGAAGACGGCCCGGCGGCCA 486

Qy         72 erLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerA 92
               |||:::     ||||||||||||   :::::::|||      ||||||   :
Db        485 AGGGTGGCCTGCAGGTGGGCGATGTGATCCTCAGCCTGAACGGCCAGTCGATCAACGAGT 426

Qy         92 laThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal---- 110
              ::     :::        :::           |||||||||   |||::::::
Db        425 CCGCCGACCTGCCGCACCTGGTGGGCAACATGAAGCCGGGCGACAAGATCAACCTGGACG 366

Qy        111 --------ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 124
                      |||  ||||||           |||::::::
Db        365 TGATTCGCAACGGCCAGCGCAAGTCCCTGAGCATGGCGGTAGGCAGCCTT 316

RESULT 11
```

```
AZ933900
LOCUS       AZ933900                719 bp    DNA     linear   GSS 24-APR-2001
DEFINITION  BJ__Ba0001B03r B. japonicum BAC library Bradyrhizobium japonicum
            genomic, genomic survey sequence.
ACCESSION   AZ933900
VERSION     AZ933900.1  GI:13775960
KEYWORDS    GSS.
SOURCE      Bradyrhizobium japonicum
  ORGANISM  Bradyrhizobium japonicum
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE   1  (bases 1 to 719)
  AUTHORS   Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A.,
            Goicoechea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.
  TITLE     A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
            genome
  JOURNAL   Genome Res. 11 (8), 1434-1440 (2001)
  MEDLINE   21376150
   PUBMED   11483585
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Class: BAC ends
            High quality sequence stop: 671.
FEATURES             Location/Qualifiers
     source          1. .719
                     /organism="Bradyrhizobium japonicum"
                     /mol_type="genomic DNA"
                     /strain="USDA110"
                     /db_xref="taxon:375"
                     /lab_host="E. coli"
                     /clone_lib="B. japonicum BAC library"
                     /note="Vector: pIndigo536; Site_1: HindIII"
ORIGIN

Alignment Scores:
Pred. No.:              4.42          Length:         719
Score:                  103.50        Matches:        34
Percent Similarity:     37.86%        Conservative:   19
Best Local Similarity:  24.29%        Mismatches:     62
Query Match:            15.33%        Indels:         25
DB:                     28            Gaps:           4

US-09-684-215B-4 (1-132) x AZ933900 (1-719)

Qy         13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArgSer 32
              |||  ||||||||||||   :::                 :::             :::
Db         44 GGCATCGGCTTTGCGATCCTTGTCAACATGGAGCGCGTGGTCGTCGCTTCCGCCAAGGGC 103

Qy         33 GlyGly----------------------GlySerProThrValHisIleGlyProThr 44
              ||||||                    |||:::          :::     |||
Db        104 GGCGGCAAGGCGGTGAAGCGTCCCTGGCTCGGCGCGAAGTTGCAGGCGGTGACGCCCGAG 163

Qy         45 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 64
                  ||||||:::         :::      ||||||  |||      |||
Db        164 ATCGCCGAAAGCCTCGGCTTG------CGTTCGCCGACCGGCGCGCTGGTCGCGAGCGTG 217

Qy         65 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 84
              |||   :::   |||||||||      |||:::   :::   |||:::||||||   :::
Db        218 GTCTCGAACGGTCCCGCAGCGAAGGCCGGCCTGAAATCCTCCGATCTGATCACCGGGATC 277

Qy         85 AspGlyAlaProIleAsn---------------------------------------- 90
              |||||      ::::::
Db        278 GACGGCCAGACCGTGGATGATCCCAATGCCTTCGACTACCGCTTCGCGACGCGTCCACTC 337

Qy         91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110
                 :::|||   |||      |||      |||          :::::::::
Db        338 GGCGGTTCCGCGCAGATCGACGGGCAGCGCGGCGGCAAGCCGCTCAAG---CTCGCGATC 394
```

```
Qy        111 ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGluGlyPro 130
              :::|||           |||     |||     |||     |||
Db        395 GCGCTGGAGACCGCACCCGACGCCGGCCGCAACGAGCTCGTCATCACTGCGCGTTCGCCG 454

RESULT 12
BZ549047/c
LOCUS       BZ549047                947 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  pacs1-60_1644.s1 pacs1-60 Pseudomonas aeruginosa genomic clone
            pacs1-60_1644, genomic survey sequence.
ACCESSION   BZ549047
VERSION     BZ549047.1  GI:27152628
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 947)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .947
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="1-60"
                     /db_xref="taxon:287"
                     /clone="pacs1-60_1644"
                     /clone_lib="pacs1-60"
                     /note="clinical isolate 1-60 Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:              6.57          Length:         947
Score:                  103.50        Matches:        37
Percent Similarity:     44.74%        Conservative:   14
Best Local Similarity:  32.46%        Mismatches:     48
Query Match:            15.33%        Indels:         16
DB:                     28            Gaps:           3

US-09-684-215B-4 (1-132) x BZ549047 (1-947)

Qy          4 SerAspAsnPheGlnLeuSerGlnGly------GlyGlnGlyPheAlaIleProIleGly 21
              |||     |||     |||  |||     |||     |||||||||||||||
Db        374 TCGCAGATCTTCACCCGTTCCGGCGGCTTCATGGGCCTGTCCTTCGCCATTCCGATCGAT 315

Qy         22 GlnAlaMetAlaIleAlaGlyGlnIleArgSer---------------------------- 32
              ||||||  :::|||  |||::::::
Db        314 GTCGCGATGAACGTCGCCGACCAGTTGAAGAAAGCCGGCAAGGTCAGTCGCGGCTGGCTG 255

Qy         33 GlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValVal 52
              |||  |||  |||                  ||||||
Db        254 GGCGT-GGTGATCCAGGAAGTGAACAACGATCTCGCCGAGTCTTCGGCCTC--------- 205

Qy         53 AspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSer 72
                 :::     :::|||||  |||  :::::|||        |||||||||
Db        204 ---GACAAGCCGTCCGGCGCGCTGGTGGCGCAACTGGTGGAAGACGGCCCGGCGGCGAAG 148

Qy         73 LeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92
              |||:::     ||||||||||||  :::::::::|||     ||||||  :::
Db        147 GGTGGCCTGCAGGTGGGCGATGTGATCAACAGCCTGAACGGCCAGTCGATCAACGAGTCC 88

Qy         93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAsp 106
```

```
                  :::       :::          |||||||||
Db        87 GCCGACCTGCCGCACCTGGTGTGGAACATGAAGCCGGGCGAC 46

RESULT 13
AQ989479
LOCUS       AQ989479                726 bp    DNA     linear   GSS 14-AUG-2000
DEFINITION  Rfc00025 Photorhabdus luminescens strain W14 M13 library
            Photorhabdus luminescens genomic clone PLG00025, genomic survey
            sequence.
ACCESSION   AQ989479
VERSION     AQ989479.1  GI:9648073
KEYWORDS    GSS.
SOURCE      Photorhabdus luminescens
  ORGANISM  Photorhabdus luminescens
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Photorhabdus.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   ffrench-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,
            Daborn,P.J., Bowen,D. and Blattner,F.R.
  TITLE     A genomic sample sequence of the entomopathogenic bacterium
            Photorhabdus luminescens W14: potential implications for virulence
  JOURNAL   Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
  MEDLINE   20378633
   PUBMED   10919786
COMMENT     Contact: ffrench-Constant RH
            Department of Biology and Biochemistry
            University of Bath
            South Building, Bath BA2 7AY, UK
            Tel: (44) 1225 826621
            Fax: (44) 1225 826779
            Email: bssrfc@bath.ac.uk
            This is one of 2,122 random reads from the M13 library. For
            annotation of identified clones (BLASTX, BLASTN and mapping to E.
            coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic
            Acids Res.
            Seq primer: M13 Forward
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Photorhabdus luminescens"
                     /mol_type="genomic DNA"
                     /strain="W14"
                     /db_xref="taxon:29488"
                     /clone="PLG00025"
                     /dev_stage="primary phase variant"
                     /clone_lib="Photorhabdus luminescens strain W14 M13
                     library"
                     /note="Genomic DNA from strain W14 was size selected (1-2
                     kb) and then cloned into M13 Janus."
ORIGIN

Alignment Scores:
Pred. No.:              5.02          Length:         726
Score:                  103.00        Matches:        39
Percent Similarity:     46.55%        Conservative:   15
Best Local Similarity:  33.62%        Mismatches:     52
Query Match:            15.26%        Indels:         11
DB:                     28            Gaps:           2

US-09-684-215B-4 (1-132) x AQ989479 (1-726)

Qy          6 AsnPheGlnLeuSerGlnGlyGlyGln--------------GlyPheAlaIleProIle 20
              :::|||     |||:::  |||:::              |||||||||||||
Db        179 TCATTCGATAAAAGTGAAAATGGTGAAACGCCAGAAGGGTTGGGGTTCGCTATTCCGACT 238

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
                 |||     |||     ::::::        |||
Db        239 GAACTCGCTACTAAGATTATGCAGAAACTTATCCGTGACGGACGAGTTATCCGGGGTTTT 298

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsn--------- 57
              ||||||  |||:::           :::     :::||||||||||
Db        299 ATCGGCATAACATCGCAAGAGCTGCCACATATTCGCTCTTCTAACGGCAATATTAACCAG 358
```

```
Qy          58 ------GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIle 75
                   |||   ||||||   :::|||       :::   ||||||          ||||||
Db         359 CTTCAAGGGCTACGAGTATTTCAGGTAACTACAAACGGGCCTGCGCAAAAAGCCGGTATT 418

Qy          76 SerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMet 95
                   ||||||:::||||||:::|||:::       |||       ||||||
Db         419 AAAGTGGGTGATATCATTACCAGCGTCAATAACAAACCCGCTATTTCTGCGGGTGAAACA 478

Qy          96 AlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThr 111
                 |||   :::       :::   ||||||   |||       ||||||
Db         479 ATGGATCAAGTGGCTGA-AATACCCCCGGCAGCGTGGGTNCTGTCACT 525

RESULT 14
AJ558965
LOCUS       AJ558965                765 bp    mRNA    linear   EST 12-JUN-2003
DEFINITION  AJ558965 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
            018_1_09_a08, mRNA sequence.
ACCESSION   AJ558965
VERSION     AJ558965.1  GI:31661537
KEYWORDS    EST.
SOURCE      Antirrhinum majus (snapdragon)
  ORGANISM  Antirrhinum majus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Lamiales; Antirrhinaceae; Antirrhineae;
            Antirrhinum.
REFERENCE   1  (bases 1 to 765)
  AUTHORS   Zachgo,S., Stueber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.
  TITLE     Antirrhinum EST collection
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Schwarz-Sommer Z
            Molekulare Pflanzengenetik
            MPI fuer Zuechtungsforschung
            Carl-von-Linne Weg 10, D-50829, Germany.
FEATURES             Location/Qualifiers
     source          1. .765
                     /organism="Antirrhinum majus"
                     /mol_type="mRNA"
                     /db_xref="taxon:4151"
                     /clone="018_1_09_a08"
                     /tissue_type="whole plant"
                     /clone_lib="Antirrhinum majus whole plant"
ORIGIN

Alignment Scores:
Pred. No.:              6.79          Length:         765
Score:                  102.00        Matches:        34
Percent Similarity:     40.83%        Conservative:   15
Best Local Similarity:  28.33%        Mismatches:     41
Query Match:            15.11%        Indels:         30
DB:                     9             Gaps:           4

US-09-684-215B-4 (1-132) x AJ558965 (1-765)

Qy          10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
               :::     |||   ||||||||||||||||          :::   |||     |||
Db         311 ACCTCGGCGGGTGTAGGATTTGCGATCCCTTCATCGACAGTGGTAAAGATAGTGCCTCAG 370

Qy          30 IleArgSerGlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGlyLeu 49
               :::             ::::::   |||                 ||||||
Db         371 TTG----------------------ATCAATTATGGAAAAGTTGTTCGAGCTGGTTTA 406

Qy          50 GlyValValAsp----------------------------AsnAsnGlyAsnGlyAla 59
                   ||||||                          |||     |||||||||
Db         407 AAT---GTTGACATAGCTCCTGATCTTATTGCAAATCAACTCAATGTTCGAAATGGAGCT 463

Qy          60 ArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGly--------------- 74
                  :::   :::|||   |||::::::   ||||||      |||
Db         464 CTAATCTTGCAGGTCCCTGGAAATAGTTCAGCAGCTAAGGCTGGGCTTGTACCTACAACA 523

Qy          75 -----------------IleSerThrGlyAspValIleThrAlaValAspGlyAlaPro 88
                              |||       |||||||:::|||   |||||||||      |||
Db         524 CGGGGTTTCGCTGGTAAAATCGTACTTGGGGATATAATTTTGGCCGTAGATGACAAACCT 583

Qy          89 IleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIle 108
               :::   |||           :::          |||:::   :::      ||||||   :::
Db         584 GTCAGGAGTAGGGCCGAGCTATACAAAACACTGGATGAGTATTCCATCGGAGACAAAGTG 643

RESULT 15
CA480614/c
LOCUS       CA480614                859 bp    mRNA    linear   EST 13-NOV-2002
DEFINITION  AGENCOURT_10739220 NCI_CGAP_ZEmb2 Danio rerio cDNA clone
            IMAGE:6789351 5', mRNA sequence.
ACCESSION   CA480614
VERSION     CA480614.1  GI:24938284
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 859)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Leonard I. Zon, M.D.
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14289  row: k  column: 14
            High quality sequence stop: 662.
FEATURES             Location/Qualifiers
     source          1. .859
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:6789351"
                     /tissue_type="embryo"
                     /lab_host="DH10B (T1-resistant)"
                     /clone_lib="NCI_CGAP_ZEmb2"
                     /note="Vector: pCMV-SPORT6.1; Site_1: EcoRV; Site_2: NotI;
                     Cloned unidirectionally. Primer: Oligo dT. Average insert
                     size 2 kb. Constructed by J. Wang (Research Genetics,
                     Invitrogen Corp) from tissue donated by L. Zon (Harvard
                     University). Note: this is a NCI_CGAP Library."
ORIGIN

Alignment Scores:
Pred. No.:              11.3          Length:         859
Score:                  100.50        Matches:        36
Percent Similarity:     40.30%        Conservative:   18
Best Local Similarity:  26.87%        Mismatches:     55
Query Match:            14.89%        Indels:         25
DB:                     14            Gaps:           5

US-09-684-215B-4 (1-132) x CA480614 (1-859)

Qy          10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIle--------- 26
               :::   |||         |||:::::::|||   |||       :::::::::
Db         854 GCCTCTGGGCACCTGAACTTCTCCTTGCCCAATGGTGATGACTTGTCCGTCAGGCAGCTC 795

Qy          27 AlaGlyGlnIleArgSerGlyGlyGlySerProThrValHisIle--------------- 41
                      ::::::   ||||||||||||          |||:::
Db         794 GTAGCTCTTCTCCAGGGAGGAGGAGGAAGCAGCGGTGCCCATCTCCTGCTCAAAGTCAAG 735

Qy          42 GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgVal 61
               |||       |||   |||   |||   |||   ||||||
Db         734 GGCCACATAGCAGAGCTTCTCCTTGATGTCACGGACAAT-----------------TTC 693

Qy          62 GlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIle 81
                        |||             ||||||||||||           :::|||:::
```

```
Db        692 CCTCTCAGCTGTGGTGGTGAAGCTGTAGCCTCTCTCGGTCAGGATCTTCATGAGGTAGTC 633

Qy         82 ThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAla--------- 98
              :::   |||      :::   :::   :::         |||
Db        632 AGTCAGGTCACGGCCAGCCAAGTCCAGACGGAGGATGGCATGGGGCAGGGCGTAACCCTC 573

Qy         99 LeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGly 118
              ::::::||||||      ||||||   |||:::                  ||||||
Db        572 GTAGATGGGCACAGTGTGGGTGACACCATCACCAGAGTCCATCACGATACCAGTGGTACG 513

Qy        119 ThrArg-----------------------ThrGlyAsnVal 124
              ||||||                      ||||||||
Db        512 ACCAGAGGCATACAGGGACAGCACAGCCTGGATGGCAACGTA 471
```

```
RESULT 16
AI597611
LOCUS       AI597611                423 bp    mRNA    linear   EST 21-APR-1999
DEFINITION  tn15f02.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2167707 3'
            similar to TR:O08664 O08664 BCL7C MRNA. ;, mRNA sequence.
ACCESSION   AI597611
VERSION     AI597611.1  GI:4606659
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .423
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2167707"
                     /tissue_type="anaplastic oligodendroglioma"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Brn25"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
ORIGIN
```

```
Alignment Scores:
Pred. No.:              5.7          Length:         423
Score:                  99.00        Matches:        35
Percent Similarity:     44.35%       Conservative:   20
Best Local Similarity:  28.23%       Mismatches:     43
Query Match:            14.67%       Indels:         26
DB:                     9            Gaps:           7

US-09-684-215B-4 (1-132) x AI597611 (1-423)
```

```
Qy          7 PheGlnLeuSerGln---------GlyGlyGlnGlyPheAlaIleProIleGlyGlnAla 23
              |||:::   ||||||          |||                 |||:::   |||
Db         39 TTTGAAAAAAGCCAAAGGGGCCCCTGGGGCAACAGGACAGGCAGGCCGGTTTCTCAGGGG 98

Qy         24 MetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThr---ValHisIleGly 42
                    ||||||:::|||     ||||||:::|||:::         :::
Db         99 TCAGGGGCATTTGGGCAGATGCGCTTGAGTGGGGGGGCACCCTCCAAGTCCTTTGTGTCA 158

Qy         43 ProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGln 62
              |||   |||      |||   |||                 |||   |||::::::
Db        159 CCCTGGGCTGCCTCAGGGACAGGT--------------GGCACTGGCTCAAAAACAGGG 203

Qy         63 ArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThr 82
                    |||::::::|||::::::|||   |||       ||||||      :::
Db        204 TAAGCTTCGGGGGCCTCACCTTCCAGCAGTTCTGGA------ACAGGCTCCTCCTTGGTC 257

Qy         83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102
              ::::::  |||:::   :::       |||
Db        258 AGCATTGGGGGTTCGTCGGTGCTGCCAGCAGTTATG----------------------- 293

Qy        103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGly------------Gly 118
                 ||||||          |||:::   |||   |||:::   |||             |||
Db        294 CCCCCGGGA---------TCTCTTTTTGGCCCACCCGTGGGGGCTGAGCCTTCTCAGGG 344

Qy        119 ThrArgThrGly 122
              |||   :::|||
Db        345 ACCCCTTCTGGG 356
```

```
RESULT 17
BI350520/c
LOCUS       BI350520                499 bp    mRNA    linear   EST 26-JUL-2002
DEFINITION  fr32b05.y1 zebrafish adult brain Danio rerio cDNA clone
            IMAGE:4955073 5' similar to SW:ACT2_FUGRU P53485 ACTIN, CYTOPLASMIC
            2 ;, mRNA sequence.
ACCESSION   BI350520
VERSION     BI350520.1  GI:15044966
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 499)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
            Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: John Ngai. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Seq primer: T7
            High quality sequence stop: 301.
FEATURES             Location/Qualifiers
     source          1. .499
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
```

/clone="IMAGE:4955073"
/sex="mixed male and female"
/tissue_type="brain"
/dev_stage="adult"
/lab_host="E. coli DH10B"
/clone_lib="zebrafish adult brain"
/note="Vector: pZIPLOX; Site_1: NotI; Site_2: SalI;
Original library was constructed in lambdaZIPLOX. Mass
excision of the cDNA library was performed to yield
pZIPLOX plasmids. Insert check was done in original
library."

ORIGIN

Alignment Scores:

Pred. No.:	10.2	Length:	499
Score:	97.50	Matches:	38
Percent Similarity:	39.29%	Conservative:	17
Best Local Similarity:	27.14%	Mismatches:	52
Query Match:	14.44%	Indels:	33
DB:	12	Gaps:	6

US-09-684-215B-4 (1-132) x BI350520 (1-499)

```
Qy          9 LeuSerGlnGly----------------------GlyGlnGlyPheAlaIleProIle 20
              |||  ||||||                      |||      :::  :::
Db        430 CTGGAACAGGGCCTCTGGACACCTGAACCTTTCATTGCCAATGGTGATGACCTGTCCGTC 371

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
                ||| :::|||:::        ::::::  ||||||||||||         |||
Db        370 AGGCAGCTGGTAGCTCTT---------CTCCAGGGAGGAGGAGGAAGCAGCGGTGCCCAT 320

Qy         41 Ile---------------GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsn 55
              :::              |||    |||  |||  |||  |||  ||||||
Db        319 CTCCTGCTCAAAGTCAAGGGCCACATAGCAGAGCTTCTCCTTGATGTCACGGACAAT--- 263

Qy         56 GlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIle 75
                                    |||                ||||||||||||
Db        262 ---------------TTCCCTCTCAGCTGTGGTGGTGAAGCTGTAGCCTCTCTCGGTCAG 218

Qy         76 SerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMet 95
                 :::|||::::::   |||      :::  :::  :::          |||
Db        217 GATCTTCATGAGGTAGTCAGTCAGGTCACGGCCAGCCAAGTCCAGACGGAGGATGGCATG 158

Qy         96 AlaAspAla---------LeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 112
                             ::::::|||||      ||||||   |||::::::
Db        157 GGGCAGGGCGTAACCCTCGTAGATGGGCACAGTGTGGGTGACACCATCACCATAGTCCAT 98

Qy        113 GlnThrLysSerGlyGlyThrArg----------------------ThrGlyAsnVal 124
                |||||| ||||||                              |||||||||
Db         97 CACGATACCAGTGGTACGACCAGAGGCATACAGGAACAGCACAGCCTGGATGGCAACGTG 38
```

RESULT 18
CA473840/c
LOCUS CA473840 870 bp mRNA linear EST 12-NOV-2002
DEFINITION AGENCOURT_10740457 NCI_CGAP_ZKid1 Danio rerio cDNA clone IMAGE:6794807 5', mRNA sequence.
ACCESSION CA473840
VERSION CA473840.1 GI:24930192
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 870)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM14303 row: n column: 22
High quality sequence stop: 671.

FEATURES Location/Qualifiers
source 1. .870
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6794807"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI_CGAP_ZKid1"
/note="Organ: kidney; Vector: pCMV-SPORT6.1; Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Constructed by J. Wang (Research Genetics, Invitrogen Corp) from tissue donated by L. Zon (Harvard University). Note: this is a NCI_CGAP Library."

ORIGIN

Alignment Scores:

Pred. No.:	22.6	Length:	870
Score:	97.50	Matches:	35
Percent Similarity:	40.62%	Conservative:	17
Best Local Similarity:	27.34%	Mismatches:	51
Query Match:	14.44%	Indels:	25
DB:	14	Gaps:	5

US-09-684-215B-4 (1-132) x CA473840 (1-870)

```
Qy         16 PheAlaIleProIleGlyGlnAlaMetAlaIle---------AlaGlyGlnIleArgSer 32
              |||::::::|||   |||      :::::::::          ::::::
Db        811 TTTTCATTGCCCAATGGTGATGACCTGTCCGTCAGGCAGCTCGTAGCTCTTCTCCAGGGA 752

Qy         33 GlyGlyGlySerProThrValHisIle---------------GlyProThrAlaPheLeu 47
              ||||||||||||          |||:::                |||     |||  |||
Db        751 GGAGGAGGAAGCAGCGGTGCCCATCTCCTGCTCAAAGTCAAGGGCCACATAGCAGAGCTT 692

Qy         48 GlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySer 67
                 |||  |||  ||||||                                    |||
Db        691 CTCCTTGATGTCACGGACAAT-----------------TTCCCTCTCAGCTGTGGTGGT 650

Qy         68 AlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAla 87
                       ||||||||||||           :::|||::::::   |||      :::
Db        649 GAAGCTGTAGCCTCTCTCGGTCAGGATCTTCATGAGGTAGTCAGTCAGGTCACGGCCAGC 590

Qy         88 ProIleAsnSerAlaThrAlaMetAlaAspAla---------LeuAsnGlyHisHisPro 104
                 :::  :::        |||                  ::::::||||||
Db        589 CAAGTCCAGACGGAGGATGGCATGGGGCAGGGCGTAACCCTCGTAGATGGGCACAGTGTG 530

Qy        105 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg------------ 120
              ||||||  |||:::          ||||||  ||||||
Db        529 GGTGACACCATCACCAGAGTCCATCACGATACCAGTGGTACGACCAGAGGCATACAGGGA 470

Qy        121 ------------ThrGlyAsnVal 124
                         |||||||||
Db        469 CAGCACAGCCTGGATGGCAACGTA 446
```

RESULT 19
CA148171
LOCUS CA148171 444 bp mRNA linear EST 24-SEP-2003
DEFINITION SCEZRZ1017F10.g RZ1 Saccharum officinarum cDNA clone SCEZRZ1017F10 5', mRNA sequence.
ACCESSION CA148171
VERSION CA148171.1 GI:35048941
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.

```
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 017  row: F  column: 10
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCEZRZ1017F10"
                     /lab_host="DH10B"
                     /clone_lib="RZ1"
                     /note="Organ: Shoot-root transition zone from young plants
                     (large insert library); Vector: pSport1; Site_1: SalI;
                     Site_2: NotI; An unidirectional cDNA library generated
                     from [Shoot-root transition zone  from young plants (large
                     insert library)]. cDNA was prepared from polyA+  mRNA
                     using SuperScript Plasmid System Kit (Invitrogen). The
                     double-strand  cDNAs were fractionated in a sepharose
                     CL-2B 40cm-columns and fragments  sizing between 0.8 and
                     1.5 Kb were directionally cloned into the vector.  Details
                     of each source of RNA and library construction can be
                     obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

Alignment Scores:
Pred. No.:                9.61          Length:          444
Score:                    97.00         Matches:         26
Percent Similarity:       47.62%        Conservative:    14
Best Local Similarity:    30.95%        Mismatches:      26
Query Match:              14.37%        Indels:          18
DB:                       13            Gaps:            2

US-09-684-215B-4 (1-132) x CA148171 (1-444)

Qy          54 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 73
               |||      |||||||||   :::   :::|||   |||   :::   ||||||
Db          32 AATGTTCGCAACGGAGCTCTTATACTTAAGGTACCTGGGGGCAGTGCTGCAGCCAAGGCA 91

Qy          74 GlyIleSer--------------------------------ThrGlyAspValIleThr 82
               |||::::::                                ||||||||||||
Db          92 GGCCTTGCTCCGACCGGCAGGGGTTTCGCTGGTAACATTGTTCTTGGTGATGTCATCGTT 151

Qy          83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102
               ||||||||||||   |||:::        :::   :::         |||:::   :::
Db         152 GCAGTGGATGGCAAACCTGTTAAGGGCAAATCAGACCTGCTGAGGGTTCTGGATGACTAT 211

Qy         103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 122
                   ||||||   ::::::::::|||                        |||   |||
Db         212 GGCGTCGGGGATCAGGTGACCTTGACA--------------------ATCCGGCGAGGC 250

Qy         123 AsnValThrLeu 126
               :::   ||||||
Db         251 TCTGAAACCCTC 262

RESULT 20
BI721127
LOCUS       BI721127                543 bp    mRNA    linear   EST 19-SEP-2001
DEFINITION  1031054B10.y1 C. reinhardtii CC-1690, Stress II (normalized),
            Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION   BI721127
VERSION     BI721127.1  GI:15696822
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 543)
  AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
            Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1031
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690, Stress II
                     (normalized), Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; Stress condition II library, constructed by John
                     Davies and Jeffrey McDermott, combines cDNAs from CC-1690
                     cells grown to mid-log phase in TAP (NH4+ - containing)
                     and shifted to TAP - NO3- (24hrs); H2 production
                     conditions (0, 12hr, 24hr) see Melis et al.,(2000) Plant
                     Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
                     sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
                     PolyA mRNA was purified from each sample, pooled and cDNA
                     synthesized. The cDNA was directionally cloned into lambda
                     Zap II (Stratagene) in the EcoRI (5') and XhoRI (3')
                     sites. pBluescript II SK- plasmids were excised from the
                     lambda ZAP clones by superinfection with ExAssist
                     (Stratagene) phage. The library was normalized using
                     method 4 described in Bonaldo et al., (1996) Genome
                     Research 6: 791-806."
ORIGIN

Alignment Scores:
Pred. No.:                12.8          Length:          543
Score:                    97.00         Matches:         31
Percent Similarity:       39.83%        Conservative:    16
Best Local Similarity:    26.27%        Mismatches:      49
Query Match:              14.37%        Indels:          22
DB:                       12            Gaps:            4

US-09-684-215B-4 (1-132) x BI721127 (1-543)

Qy          10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
               :::      |||   |||||||||:::||||||                :::     |||
Db         153 GCATCCTCCGGCGTGGGCTTCGCGCTGCCCATCGACAGTGTGCGCGGCCTGGTAGACCAG 212

Qy          30 IleArgSerGlyGlyGly------SerProThrValHisIleGly---------ProThr 44
               |||   :::   |||              |||   :::   :::           |||
Db         213 ATCCTGACCTACGGCCGCGTGCTGCGGCCCGTGCTGGGAGTCAGCCTGGCGCCGCCGCAG 272

Qy          45 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 64
                     ||||||                        |||     |||     |||
Db         273 GTGCTCAAGCAGCTGGGC--------------CAGCCGGGCGTGCTGGTGCTAGAGGTG 317

Qy          65 ValGlySerAlaProAlaAlaSerLeuGlyIleSer---------------------- 76
                     :::|||||||        ||||||
Db         318 CCCAAGGGCTCCCCCGCGGAGAAGGCCGGCATCAAGCCCACCATGCGCGACCGCTTCAGC 377

Qy          77 ------------ThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92
```

```
Db        378 GGCTCCCTGGTGCTGGGCGACATCATCACAGGCATTGACGGCAAGGCCGTAAAGAACTAT 437

Qy         93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110
Db        438 TCGGACCTGGTGGAGGCGCTGGATGAGAAGCGCGTGGGCGACACAGTCAAGGTG 491

RESULT 21
AW285510
LOCUS       AW285510                546 bp    mRNA    linear   EST 19-JUL-2000
DEFINITION  LG1_241_E05.g1_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   AW285510
VERSION     AW285510.1  GI:6675354
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 546)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
  TITLE     An EST database from Sorghum: light-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: T7
            High quality sequence start: 126
            High quality sequence stop: 546
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .546
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /db_xref="taxon:4558"
                     /clone_lib="Light Grown 1 (LG1)"
                     /note="Organ: 10- to 14-day-old light-grown (greenhouse)
                     seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2:
                     EcoRI; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
ORIGIN

Alignment Scores:
Pred. No.:                16.2          Length:          546
Score:                    96.00         Matches:         25
Percent Similarity:       47.62%        Conservative:    15
Best Local Similarity:    29.76%        Mismatches:      26
Query Match:              14.22%        Indels:          18
DB:                       10            Gaps:            2

US-09-684-215B-4 (1-132) x AW285510 (1-546)

Qy         54 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 73
Db         56 AATGTTCGCAACGGAGCTCTTATACTTAAGGTACCTGGGGGCAGTGCTGCAGCCAAAGCA 115

Qy         74 GlyIleSer--------------------------------ThrGlyAspValIleThr 82
Db        116 GGTCTTGCTCCAACCGGCAGGGGTTTTGCTGGTAATATTGTTCTGGGTGATATCATCGTT 175

Qy         83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102
Db        176 GCAGTGGACGGCAAACCTGTTAAGGGCAAATCTGACCTGCTGAGGGTTCTGGATGACTAT 235

Qy        103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 122
Db        236 GGCGTCGGAGATCAGGTGACCTTGACA--------------------ATCCGGCGAGGC 274

Qy        123 AsnValThrLeu 126
Db        275 TCAGAAACCCTT 286

RESULT 22
AW285527
LOCUS       AW285527                551 bp    mRNA    linear   EST 19-JUL-2000
DEFINITION  LG1_241_G05.g1_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   AW285527
VERSION     AW285527.1  GI:6675371
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 551)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
  TITLE     An EST database from Sorghum: light-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: T7
            High quality sequence start: 16
            High quality sequence stop: 552
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .551
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /db_xref="taxon:4558"
                     /clone_lib="Light Grown 1 (LG1)"
                     /note="Organ: 10- to 14-day-old light-grown (greenhouse)
                     seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2:
                     EcoRI; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
ORIGIN

Alignment Scores:
Pred. No.:                16.5          Length:          551
Score:                    96.00         Matches:         25
Percent Similarity:       47.62%        Conservative:    15
Best Local Similarity:    29.76%        Mismatches:      26
Query Match:              14.22%        Indels:          18
DB:                       10            Gaps:            2

US-09-684-215B-4 (1-132) x AW285527 (1-551)

Qy         54 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 73
Db         61 AATGTTCGCAACGGAGCTCTTATACTTAAGGTACCTGGGGGCAGTGCTGCAGCCAAAGCA 120

Qy         74 GlyIleSer--------------------------------ThrGlyAspValIleThr 82
Db        121 GGTCTTGCTCCAACCGGCAGGGGTTTTGCTGGTAATATTGTTCTGGGTGATATCATCGTT 180

Qy         83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102
Db        181 GCAGTGGACGGCAAACCTGTTAAGGGCAAATCTGACCTGCTGAGGGTTCTGGATGACTAT 240
```

```
Qy        103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 122
              ||||||   ::::::::::|||                          |||   |||
Db        241 GGCGTCGGAGATCAGGTGACCTTGACA--------------------ATCCGGCGAGGC 279

Qy        123 AsnValThrLeu 126
              :::   ||||||
Db        280 TCAGAAACCCTT 291

RESULT 23
BU557763/c
LOCUS       BU557763               1093 bp    mRNA    linear   EST 16-SEP-2002
DEFINITION  AGENCOURT_10214312 NIH_MGC_107 Homo sapiens cDNA clone
            IMAGE:6585799 5', mRNA sequence.
ACCESSION   BU557763
VERSION     BU557763.1  GI:22908059
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1093)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2798  row: f  column: 07
            High quality sequence stop: 299.
FEATURES             Location/Qualifiers
     source          1. .1093
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6585799"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_107"
                     /note="Organ: breast; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              44.2          Length:        1093
Score:                  96.00         Matches:       48
Percent Similarity:     35.43%        Conservative:  14
Best Local Similarity:  27.43%        Mismatches:    46
Query Match:            14.22%        Indels:        68
DB:                     13            Gaps:          9

US-09-684-215B-4 (1-132) x BU557763 (1-1093)

Qy         23 AlaMetAlaIleAlaGlyGlnIleArgSerGlyGly------------------GlySer 36
              |||   |||         |||   |||:::||||||
Db        789 GCTCGGGCGCATGGGAGTCAGGCCAGAAACGGGGGCTGGATGGCGGGGCGGCTAAGCTAT 730

Qy         37 ProThrValHisIleGlyProThrAla-------------------------------- 45
              ||||||         |||||||||
Db        729 CCCACGGCGCCCCGCGGCCCCACGGTGCAGAGCCCCCAGGGCGGCCGGGGCACGCCTCGG 670

Qy         46 ------------------PheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAla 59
                               |||   |||:::         ||||||:::||||||
Db        669 GCCAAGGGGGGCACCTGGAGTAGGGGGTACGGGATTTGC---GGCAATGGCTCAGGCGCT 613

Qy         60 ArgValGln----------------------------------ArgValValGly--- 66
              |||:::|||                                 |||:::|||
Db        612 CGCCTACAACCCGTAAAACCAGTCTCCGGTGGAGGATCTCTCGGGGAGGTAATTGGGTGC 553

Qy         67 ---------SerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAla 83
                          |||   |||       |||::::::|||||||||   :::
Db        552 AGGGACGCATGCCAACCAGGGGCCGGGCGCGGAGTGGCGACGGGGGACGGCCTACACCTG 493

Qy         84 ValAsp------------GlyAlaProIleAsnSerAlaThrAlaMetAlaAspAla--- 98
              |||:::            |||  |||           |||:::|||   ||||||
Db        492 GTAGAGTGGCACCGGGGTGGGGGACCC---CGTGTCCAGACGTCCATGCCCGATGCCAGG 436

Qy         99 --------------------------------LeuAsnGlyHis--------------- 102
                                              |||:::||||||
Db        435 GTGGCGCCCAGAGGGATAGAGGGACAGCACGGCCTGGATGGCCACGTACATGACCGGGGG 376

Qy        103 --------------HisProGlyAspValIleSerValThrTrpGlnThrLysSerGly 117
                           |||   |||         |||:::   |||          ||||||
Db        375 GTTGAAGGTGTCAAACATAATGGGAGTCA-TCTTCTCTCTGTTGGCCTTGGGGTTCAGGG 317

Qy        118 GlyThrArgThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              |||   |||:::      :::            ||||||   |||
Db        316 GGACCTCGGTCAGCAGCACTGGGTGCTCCTCCGGGGCCACGCGCA 272

RESULT 24
BZ576076
LOCUS       BZ576076               1394 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  msh2_4761.y2 msh Pseudomonas aeruginosa genomic clone msh2_4761,
            genomic survey sequence.
ACCESSION   BZ576076
VERSION     BZ576076.1  GI:27211137
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 1394)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .1394
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="MSH"
                     /db_xref="taxon:287"
                     /clone="msh2_4761"
                     /clone_lib="msh"
                     /note="Environmental isolate. Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:              62.8          Length:        1394
Score:                  96.00         Matches:       42
Percent Similarity:     37.20%        Conservative:  19
Best Local Similarity:  25.61%        Mismatches:    43
Query Match:            14.22%        Indels:        60
DB:                     28            Gaps:          9
```

US-09-684-215B-4 (1-132) x BZ576076 (1-1394)

```
Qy          12 GlyGlyGlnGlyPheAlaIleProIleGly-----------------------GlnAla 23
               ||||||||||||      |||   :::|||                         |||
Db         320 GGCGGACAGGGAGTAGGGATCGAACTCGGTGGCGCGCGCCGGGTAGTCCTTGCGCAGCGG 379

Qy          24 MetAlaIle--------------------AlaGlyGlnIleArgSerGlyGlyGlySer 36
               ||||||:::                  ||||||||||:::|||     |||
Db         380 ATGGCCCTGCCAGGTCGGCGGCATCAGCATGCGGGTCAGGTGCGGGTGGCCGGTGAAGGT 439

Qy          37 ProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGly 56
                  ||||||:::         |||   :::   :::||||||      :::   |||
Db         440 GATGCCGTACATGTCCCACACTTCGCGCTCGTACCAGTTGGCGTTGGGCCAGATACGGGT 499

Qy          57 AsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSer 76
                  |||   :::|||:::      :::|||
Db         500 GGCGGTGGGCAGGTTGAGGTCGCGCTCGGACAA-------------------------- 532

Qy          77 ThrGlyAspVal---------------IleThr---------------AlaValAspGly 86
                  ||||||:::                ||||||               |||:::
Db         533 ---GGCGACCTTGATCATCACGTCACTATTACGCTCCAGCGACATCAGGTGGTAGAACAC 589

Qy          87 AlaProIleAsnSerAlaThrAlaMetAla--------------------AspAlaLeu 99
               |||   :::   :::      |||   |||                   ||||||:::
Db         590 GCTGAAGTCGGCGCTCGGCAGGCCGCGGCGATGGGTGCGCAGCCGCTCGTCGACGCCGTG 649

Qy         100 Asn------GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGly 117
                     ||||||  |||                 |||                 |||
Db         650 CAAGTCGTAGAGCATCACGTAGGG---------------CTTTGGGTACCTGACGGAGGA 694

Qy         118 --------------------------------GlyThrArgThrGlyAsnValThrLeu 126
                                              ||||||:::::      :::   :::
Db         695 AGGGTCAGGGACTTCGATCCAGTCTTTTCCGCGGGAACCCAGAGCACCGGCATGCCGGTA 754

Qy         127 AlaGluGlyPro 130
               ||||||
Db         755 CCGGGTGGGCCT 766
```

RESULT 25
CF787130/c

```
LOCUS       CF787130                600 bp    mRNA    linear   EST 21-OCT-2003
DEFINITION  854999 MARC 4PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   CF787130
VERSION     CF787130.1  GI:37791691
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
            Nonneman,D.J., Wray,J.E. and Keele,J.W.
  TITLE     Porcine EST collection using a normalized library constructed from
            embryos representing early developmental stages
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: TMW8001  row: I  column: 11
            Seq primer: GTAATACGACTCACTATAGGG.
FEATURES             Location/Qualifiers
     source          1. .600
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 4PIG"
                     /note="Vector: pcDNA3.1; Site_1: EcoRI; Site_2: NotI;
                     Library made with combined RNA from day-10, day-13,
                     day-15, day-25, and day-30 whole embryos."
ORIGIN
```

Alignment Scores:

Pred. No.:	20.8	Length:	600
Score:	95.50	Matches:	33
Percent Similarity:	45.00%	Conservative:	21
Best Local Similarity:	27.50%	Mismatches:	45
Query Match:	14.15%	Indels:	21
DB:	14	Gaps:	6

US-09-684-215B-4 (1-132) x CF787130 (1-600)

```
Qy          12 GlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArg 31
               |||      |||          |||      |||                ||||||:::|||
Db         554 GGGGCAACAGGACAAGGCAGGCCGGCTTCTCAGGGGTCAGGGGCATTTGGGCAGATGCGC 495

Qy          32 SerGlyGlyGlySerProThrVal---HisIleGlyProThrAlaPheLeuGlyLeuGly 50
                  ||||||:::|||:::|||      :::   |||   |||       |||   |||
Db         494 TTGAGTGGGGGGGGCGCCCTCCGTGTCCTCTGTATCACCCTGGGCTGCCTCGGGGACAGGT 435

Qy          51 ValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAla 70
                     |||   |||::::::                ||||::::::   :::
Db         434 --------------GGCACTGGCTCAAAGACGGGGTAAGCTTCAGGGGCCTCAGCCTCC 390

Qy          71 AlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsn 90
               ::::::   |||      ||||||      :::   ::::::   |||:::  :::
Db         389 AGCAATTCTGGA------ACAGGCTCCTCCTTGGTCAGCATCGGGGGTTCATCCGTGCCG 336

Qy          91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110
                  |||                 |||      ||||||         |||:::
Db         335 CCTGCA----------------------GGTATGCCCCCAGGA---------TCTCTC 309

Qy         111 ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGluGlyPro 130
               :::|||   :::      |||         :::|||      |||   :::   ||||||
Db         308 TCCTGGCCCAGCTGTGGGGGCTGAGCATCCTCAGGG------ACCCCTTCCGGTGGTCCA 255
```

RESULT 26
BX674871/c

```
LOCUS       BX674871                647 bp    mRNA    linear   EST 28-OCT-2003
DEFINITION  BX674871 Sus Scrofa library (scac) Sus scrofa cDNA clone
            scac0037i.j.05 5prim, mRNA sequence.
ACCESSION   BX674871
VERSION     BX674871.1  GI:38008823
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 647)
  AUTHORS   Bonnet,A., Tosser-Klopp,G., Benne,F., Cabau,C., Villeger,S.,
            Soares,M., Bonaldo,F. and Hatey,F.
  TITLE     A Pig Normalised Multi-Tissue cDNA Library
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tosser-Klopp G
            Genetique Animale
            Institut National de la Recherche Agronomique
            Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan
            cedex, FRANCE
            Tel: 33 (0) 5.61.28.51.14
            Fax: 33 (0) 5.61.28.53.08
            Email: tosser@toulouse.inra.fr
            Clone distribution: AGENAE Resource centre. Francois PIUMI,
            Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du
            genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex,
            FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73
            Sequence cleaned of vector, adaptator and repetitions. Contact us
            at sigenasupport@jouy.inra.fr to obtain the chromatogram of this
            sequence.
```

```
           Plate: 0037,  row: j  column: 5.
FEATURES             Location/Qualifiers
     source          1. .647
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /clone="scac0037i.j.05"
                     /tissue_type="mixed"
                     /clone_lib="Sus Scrofa library (scac)"
                     /note="Vector: pT7T3D-pac vector; tissues: adipose tissue,
                     brain, kidney, liver, muscle, ovary, testis,  heart,
                     hypothalamus, pancreas, skin, spleen, thymus, placenta,
                     pituitary  gland, seminal vesicle, small intestine,
                     uterus, adrenals, bulbo uretral  gland, cerebral trunk,
                     epididymis, female gonad, gall-bladder,  hippocampus,
                     large intestine, male gonad, melanocytes, stomach, udder"
ORIGIN

Alignment Scores:
Pred. No.:                 23.2          Length:          647
Score:                     95.50         Matches:         33
Percent Similarity:        45.00%        Conservative:    21
Best Local Similarity:     27.50%        Mismatches:      45
Query Match:               14.15%        Indels:          21
DB:                        13            Gaps:            6

US-09-684-215B-4 (1-132) x BX674871 (1-647)

Qy          12 GlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArg 31
               |||     |||       |||     |||                 ||||||:::|||
Db         600 GGGGCAACAGGACAAGGCAGGCCGGCTTCTCAGGGGTCAGGGGCATTTGGGCAGATGCGC 541

Qy          32 SerGlyGlyGlySerProThrVal---HisIleGlyProThrAlaPheLeuGlyLeuGly 50
               ||||||:::|||:::|||      :::   |||   |||      |||   |||
Db         540 TTGAGTGGGGGGGCGCCCTCCGTGTCCTCTGTATCACCCTGGGCTGCCTCGGGGACAGGT 481

Qy          51 ValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAla 70
                        |||   |||::::::                |||::::::   :::
Db         480 --------------GGCACTGGCTCAAAGACGGGGTAAGCTTCAGGGGCCTCAGCCTCC 436

Qy          71 AlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsn 90
               ::::::   |||      ||||||      :::   ::::::   |||:::   :::
Db         435 AGCAATTCTGGA------ACAGGCTCCTCCTTGGTCAGCATCGGGGGTTCATCCGTGCCG 382

Qy          91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110
                  |||                 |||      |||||         |||:::
Db         381 CCTGCA----------------------GGTATGCCCCCAGGA---------TCTCTC 355

Qy         111 ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGluGlyPro 130
               :::|||   :::      |||      :::|||      |||   :::   ||||||
Db         354 TCCTGGCCCAGCTGTGGGGGCTGAGCATCCTCAGGG------ACCCCTTCCGGTGGTCCA 301

RESULT 27
BI378928
LOCUS       BI378928                677 bp    mRNA    linear   EST 26-AUG-2003
DEFINITION  BFLG1_000559 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG
            or MPMGp498) Branchiostoma floridae cDNA clone MPMGp498P1235 5',
            mRNA sequence.
ACCESSION   BI378928
VERSION     BI378928.1  GI:30914099
KEYWORDS    EST.
SOURCE      Branchiostoma floridae (Florida lancelet)
  ORGANISM  Branchiostoma floridae
            Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
            Branchiostoma.
REFERENCE   1  (bases 1 to 677)
  AUTHORS   Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J.,
            Herwig,R., Vingron,M. and Lehrach,H.
  TITLE     New evidence for genome-wide duplications at the origin of
            vertebrates using an amphioxus gene set and completed animal
            genomes
  JOURNAL   Genome Res. 13 (6A), 1056-1066 (2003)
  MEDLINE   22683279
  PUBMED    12799346
COMMENT     Contact: Panopoulou G
            laboratory 145, dept.Lehrach
            Max-Planck-Institut fuer Molekulare Genetik
            Ihnestr.63-73, D-14195 Berlin, Germany
            Tel: +49 30 8413 1235
            Fax: +49 30 8413 1128
            Email: panopoul@molgen.mpg.de
            The library was characterised by oligonucleotide fingerprinting
            (ONFP) to reduce sequencing redundancy. According to the ONFP
            procedure, clones giving the same hybridisation pattern with a
            battery of 200 8mer oligonucleotides are grouped into clusters. One
            clone per cluster is selected for sequencing. The size of each
            cluster is an indicator of the frequency of a transcript in the
            analysed library. The cluster size as well the coordinates of the
            rest of the clones assigned to the same fingerprint cluster as the
            clone from which the above EST is generated is available at the
            amphioxus project site at: http://www.molgen.mpg.de/amphioxus/
            Clones and filters are distributed via the Resource Center/Primary
            Database of the German Genome Project (http://www.rzpd.de)
            PCR PRimers
            FORWARD: 5' CCCCAGGCTTTACACTTTATGCTTCCGGCTCG 3' (M13RSP)
            BACKWARD: 5' GCTATTACGCCAGCTGGCGAAAGGGGGATGTG 3' (M13FSP)
            Insert Length: 1200   Std Error: 0.00
            Seq primer: 5'-CCGGTCCGGAATTCCCGGGT-3' pSport3/86
            High quality sequence stop: 677.
FEATURES             Location/Qualifiers
     source          1. .677
                     /organism="Branchiostoma floridae"
                     /mol_type="mRNA"
                     /db_xref="taxon:7739"
                     /clone="MPMGp498P1235"
                     /tissue_type="whole embryo"
                     /dev_stage="5-6 hrs (gastrula stage)"
                     /lab_host="E.coli, XL1 blue"
                     /clone_lib="Amphioxus 5-6 hrs cDNA library (Name
                     convention: BFLG or MPMGp498)"
                     /note="Vector: pSport1; Site_1: SalI, KpnI, EcoRI (5');
                     Site_2: NotI, BamHI, HindIII (3'); OligodT primed and
                     directionally cloned in pSport1 vector using a  NotI
                     (5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC (T)15-3' and a SalI 5'-
                     TCGACCCACGCGTCCG-3'adapters (Gibco BRL)."
ORIGIN

Alignment Scores:
Pred. No.:                 24.8          Length:          677
Score:                     95.50         Matches:         21
Percent Similarity:        61.11%        Conservative:    12
Best Local Similarity:     38.89%        Mismatches:      18
Query Match:               14.15%        Indels:          3
DB:                        12            Gaps:            1

US-09-684-215B-4 (1-132) x BI378928 (1-677)

Qy          58 GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThr 77
               |||     |||   ||:::::      :::||||||   |||   ||||||
Db         143 GGAGTACTGGTGCACAGGATCATTGTGGGGTCACCTGCATACAGTGCGGGTATCCGTCCT 202

Qy          78 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 97
               |||||||||||||:::::::::|||      :::   ||||||      :::   |||
Db         203 GGTGACGTCATCACCAGTATCAACGGTCGATAGGTGACGTCAGCACGAGACATCTACGAC 262

Qy          98 AlaLeuAsnGlyHisHisProGlyAspValIleSerValThr 111
               |||:::|||           |||      :::::::::|||
Db         263 GCGGTAAATAGC---------GGGCAGCAGCTGAACATCACC 295

RESULT 28
CK161112
LOCUS       CK161112                881 bp    mRNA    linear   EST 05-DEC-2003
DEFINITION  FGAS042794 Triticum aestivum FGAS: TaLt5 Triticum aestivum cDNA,
            mRNA sequence.
ACCESSION   CK161112
VERSION     CK161112.1  GI:38988986
```

```
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 881)
  AUTHORS   Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
            Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
            Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,
            Penniket,C., Roach,J.L. and Sarhan,F.
  TITLE     Functional Genomics of Abiotic Stress In Wheat and Canola Crops
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033
            Email: fgas_ests@cs.usask.ca
            This sequence is the direct result of the Base calling software
            Phred (default parameters). It is the raw base calls. To aid in the
            identification of the high quality insert the software Lucy
            (default parameters) has been run on this sequence. Lucy identified
            the region [123,692].
            Plate: TaLt547  row: K  column: 14.
FEATURES             Location/Qualifiers
     source          1. .881
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Wheat line PI 178383"
                     /db_xref="taxon:4565"
                     /lab_host="DH5 alpha"
                     /clone_lib="Triticum aestivum FGAS: TaLt5"
                     /note="Organ: Crown; Vector: pGEM-T; SSH (suppression
                     subtractive hybridization) cDNA library from genotype
                     PI178383 cold hardened at 2 C for 21 days and 49 days
                     (equal amount of cDNA pooled together before subtraction,
                     tester) and subtracted against genotype Norstar cold
                     hardened at 2 C for 1 day (24 H)(driver). Modified Smart
                     cDNA (Clontech)priming and non-directional cloning"
ORIGIN

Alignment Scores:
Pred. No.:              40.6          Length:        881
Score:                  95.00         Matches:       35
Percent Similarity:     44.88%        Conservative:  22
Best Local Similarity:  27.56%        Mismatches:    58
Query Match:            14.07%        Indels:        12
DB:                     14            Gaps:          6

US-09-684-215B-4 (1-132) x CK161112 (1-881)

Qy          7 PheGlnLeuSerGlnGly---GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAla 25
              |||   :::      |||   |||   ||||||   |||   :::|||
Db        509 TTTGGCATCGGCTCTGGTTTGGGCTCTGGCTTCGGCATAGGCTTGGGCTCGGGCTTTGGC 568

Qy         26 IleAlaGlyGlnIleArgSerGlyGlyGlySerPro---ThrValHisIleGlyProThr 44
              |||        :::   ||||||   ||||||        :::   :::|||   :::
Db        569 ATCGGCTCAGGCTTGGGCTCTGGCTTCGGTTCCGGCTTTGGCATTGGCTTAGGC---TCT 625

Qy         45 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 64
                 |||:::|||   |||:::          |||:::|||          :::
Db        626 GGCTTCATTGGCTCGGGTTTGGGCTCTGGCTTTGGTTCTGGTTTGGGCATCGGCTCANGC 685

Qy         65 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 84
              :::||||||         :::   |||||||||      |||      :::   :::
Db        686 TTGGGCTCANGCTTAGGCTCAGGTTTGGGCATTGGGTAAGGC------TTGGGCTCCGGC 739

Qy         85 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 104
                 |||:::              :::   :::      ||||||      |||   |||
Db        740 TTCGGTTCGGGTTTTGGCATCGGCTCAGGCTTGGGCTCTGCGTTGATC---CACTGGCCG 796

Qy        105 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 124
              |||                    ::::::||||||                |||::::::
Db        797 GGCGGG----------------CGCTCGAAATCCGCGGCCATGGCGGCCGGGAGCATG 838

Qy        125 ThrLeuAlaGluGlyProPro 131
                 :::       ||||||
Db        839 CACGTCNGGGCCATTCCGCCT 859

RESULT 29
CA975497/c
LOCUS       CA975497                909 bp    mRNA    linear   EST 06-JAN-2003
DEFINITION  AGENCOURT_11067943 NCI_CGAP_ZKid1 Danio rerio cDNA clone
            IMAGE:6805289 5', mRNA sequence.
ACCESSION   CA975497
VERSION     CA975497.1  GI:27508151
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 909)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Leonard I. Zon, M.D.
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14331  row: c  column: 16
            High quality sequence start: 10
            High quality sequence stop: 695.
FEATURES             Location/Qualifiers
     source          1. .909
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:6805289"
                     /lab_host="DH10B (T1-resistant)"
                     /clone_lib="NCI_CGAP_ZKid1"
                     /note="Organ: kidney; Vector: pCMV-SPORT6.1; Site_1:
                     EcoRV; Site_2: NotI; Cloned unidirectionally. Primer:
                     Oligo dT. Average insert size 1.8 kb. Constructed by J.
                     Wang (Research Genetics, Invitrogen Corp) from tissue
                     donated by L. Zon (Harvard University). Note: this is a
                     NCI_CGAP Library."
ORIGIN

Alignment Scores:
Pred. No.:              42.5          Length:        909
Score:                  95.00         Matches:       38
Percent Similarity:     39.86%        Conservative:  17
Best Local Similarity:  27.54%        Mismatches:    55
Query Match:            14.07%        Indels:        28
DB:                     14            Gaps:          6

US-09-684-215B-4 (1-132) x CA975497 (1-909)

Qy          9 LeuSerGlnGlyGlyGlnGly---------PheAlaIleProIleGlyGlnAlaMetAla 25
              |||   ||||||      |||         |||   :::|||   |||      ::::::
Db        808 CTGGAACAGGGGCCTCTGGGGCACTTGAACTTCTTCATGCCCAATGGTGATGACCTGTCC 749

Qy         26 Ile---------AlaGlyGlnIleArgSerGlyGlyGlySerProThrValHisIle--- 41
              :::                  ::::::   ||||||||||||          |||:::
Db        748 GTCAGGCAGCTCGTAGCTCTTCTCCAGGGAGGAGGAGGAAGCAGCGGTGCCCATCTCCTG 689

Qy         42 -----------GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsn 57
```

```
Db        688 CTCAAAGTCAAGGGCCACATAGCAGAGCTTCTCCTTGATGTCACGGACAAT--------- 638

Qy         58 GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThr 77

Db        637 ---------TTCCCTCTCAGCTGTGGTGGTGAAGCTGTAGCCTCTCTCGGTCAGGATCTT 587

Qy         78 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 97

Db        586 CATGAGGTAGTCAGTCAGGTCACGGCCAGCCAAGTCCAGACGGAGGATGGCGTGGGGCAG 527

Qy         98 Ala---------LeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThr 114

Db        526 GGCGTAACCCTCGTAGATGGGCACAGTGTGGGTGACACCATCACCAGAGTCCATCACGAT 467

Qy        115 LysSerGlyGlyThrArg----------------------ThrGlyAsnVal 124

Db        466 ACCAGTGGTACGACCAGAGGCATACAGGGACAGCACAGCCTGGATGGCAACGTA 413
```

```
RESULT 30
BJ285991/c
LOCUS       BJ285991                728 bp    mRNA    linear   EST 09-APR-2002
DEFINITION  BJ285991 Y. Ogihara unpublished cDNA library, Wh_r Triticum
            aestivum cDNA clone whr13l24 3', mRNA sequence.
ACCESSION   BJ285991
VERSION     BJ285991.1  GI:20105650
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 728)
  AUTHORS   Ogihara,Y. and Murai,K.
  TITLE     Expressed genes in Triticum aestivum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .728
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="whr13l24"
                     /tissue_type="root"
                     /dev_stage="Feekes' scale 1"
                     /clone_lib="Y. Ogihara unpublished cDNA library, Wh_r"
ORIGIN
```

```
Alignment Scores:
Pred. No.:               34.6          Length:         728
Score:                   94.50         Matches:        37
Percent Similarity:      42.22%        Conservative:   20
Best Local Similarity:   27.41%        Mismatches:     53
Query Match:             14.00%        Indels:         25
DB:                      12            Gaps:           5

US-09-684-215B-4 (1-132) x BJ285991 (1-728)

Qy         10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIle--------- 26

Db        582 ACGTCCGCCGGTGTTGGTTTTGCTATCCAATCATCAACTATACTCAAAATAGTTCCTCAG 523

Qy         27 ---------AlaGlyGlnIleArgSerGlyGlyGlySerProThrValHisIleGlyPro 43

Db        522 TTGATTCAGTCTGGAAAGGTTCGACGTGCTGGGCTGAATGTGGAATTCGCACCAGATCCG 463

Qy         44 ThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArg 63

Db        462 ATTGCGTAT---CAGCTTAACGTCCGTGAT-----------GGTGCTCTTATACTGAAG 418

Qy         64 ValValGlySerAlaProAlaAlaSerLeuGly--------------------------- 74

Db        417 GTTCCCCCGGGCAGCACCGTGGAAAAAGCAGGTCTAGTTCCTACAGGCAGGGGTTTTGCC 358

Qy         75 ------IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92

Db        357 GGCAATATTGTCCTGGGCGACGTCATTGTTGCAGTGGATGGCAAACCGATTAAGGGCAAA 298

Qy         93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 112

Db        297 TCTGACCTCCTGAGGGTTCTAGACGACTACGGGGTCGGAGACACGGTGACCCTGACGATC 238

Qy        113 GlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeuAla 127

Db        237 AGAAGA---------GGCACTGAGACAATTCCGATAGCCCTGTCG 202
```

```
Search completed: April 30, 2004, 04:40:50
Job time : 2593.37 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 20:30:12 ; Search time 375.474 Seconds
(without alignments)
1493.479 Million cell updates/sec

Title: US-09-684-215B-4
Perfect score: 675
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....QTKSGGTRTGNVTLAEGPPA 132

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061304_13179/app_query.fasta_1.1180
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_819_@runat_29042004_061304_13179 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	675	100.0	396	4	AAL40769	Aal40769 Nucleotid
2	675	100.0	672	4	AAL40772	Aal40772 Nucleotid
3	675	100.0	702	2	AAZ20206	Aaz20206 Mycobacte
4	675	100.0	702	4	AAL40770	Aal40770 Nucleotid
5	675	100.0	702	6	ABK14140	Abk14140 DNA encod
6	675	100.0	1002	6	AAD47078	Aad47078 Mycobacte
7	675	100.0	1002	6	AAD47077	Aad47077 Mycobacte
8	675	100.0	1002	6	AAD28336	Aad28336 Mycobacte
9	675	100.0	1002	6	AAD28337	Aad28337 Mycobacte
10	675	100.0	1068	2	AAX34251	Aax34251 Mycobacte
11	675	100.0	1143	2	AAX34252	Aax34252 Mycobacte
12	675	100.0	1629	5	AAS03793	Aas03793 M. tuberc
13	675	100.0	1742	4	AAL40771	Aal40771 Nucleotid
14	675	100.0	1871	4	AAL40768	Aal40768 Nucleotid
15	675	100.0	2181	8	ADA26353	Ada26353 Mycobacte
16	675	100.0	2190	6	AAD47084	Aad47084 Mycobacte
17	675	100.0	2190	6	AAD28343	Aad28343 Mycobacte
18	675	100.0	2191	4	AAL40773	Aal40773 Nucleotid
19	675	100.0	2286	6	ABK14128	Abk14128 DNA encod
20	675	100.0	2287	2	AAZ20194	Aaz20194 Mycobacte
21	675	100.0	2287	6	AAD47083	Aad47083 Mycobacte
22	675	100.0	2287	6	AAD28342	Aad28342 Mycobacte
23	675	100.0	2451	8	ADA26360	Ada26360 Mycobacte
24	675	100.0	2487	8	ADA26359	Ada26359 Mycobacte
25	675	100.0	2637	8	ADA26358	Ada26358 Mycobacte
26	675	100.0	2808	6	AAD47110	Aad47110 Mycobacte
27	675	100.0	2808	8	ADA26357	Ada26357 Mycobacte
28	675	100.0	3030	8	ADA26355	Ada26355 Mycobacte
29	675	100.0	3060	8	ADA26363	Ada26363 M. bovis
30	675	100.0	3104	8	ADA26362	Ada26362 Mycobacte
31	675	100.0	3474	8	ADA26361	Ada26361 Mycobacte
32	675	100.0	110000	4	AAI99682_01	Continuation (2 of
33	675	100.0	110000	4	AAI99683_01	Continuation (2 of
34	670	99.3	447	2	AAT91403	Aat91403 Mycobacte
35	670	99.3	447	2	AAT91466	Aat91466 Mycobacte
36	670	99.3	447	2	AAV44342	Aav44342 Mycobacte
37	670	99.3	447	2	AAV64450	Aav64450 M. tuberc
38	670	99.3	447	2	AAZ19040	Aaz19040 M. tuberc
39	670	99.3	447	2	AAZ19252	Aaz19252 M. tuberc
40	670	99.3	447	5	AAS03780	Aas03780 M. tuberc
41	670	99.3	447	6	AAD47080	Aad47080 Mycobacte
42	670	99.3	447	6	AAD28339	Aad28339 Mycobacte
43	670	99.3	1872	2	AAT91414	Aat91414 Mycobacte
44	670	99.3	1872	2	AAT91477	Aat91477 Mycobacte
45	670	99.3	1872	2	AAV44355	Aav44355 Mycobacte

ALIGNMENTS

RESULT 1
AAL40769
ID AAL40769 standard; DNA; 396 BP.
XX
AC AAL40769;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ra12 protein.
XX
KW Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1. .396
FT /*tag= a
FT /product= "Ra12 protein"
FT /note= "No start or stop codon"
XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027652.
XX
PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.
XX

```
PI   Skeiky Y,  Guderian J;
XX
DR   WPI; 2001-266299/27.
DR   P-PSDB; AAO22138.
XX
PT   Recombinant nucleic acid molecule for producing high yield expression of
PT   desired fusion polypeptides, encodes fusion polypeptide comprising
PT   Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS   Claim 1; Fig 2; 39pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC   kDa C-terminal fragment of serine protease antigen MTB32A of
CC   Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC   The recombinant fusion nucleic acids and polypeptides are useful for
CC   providing stable and high yield expression of fusion polypeptides of both
CC   eukaryotic and prokaryotic origin and to encode a protein product for use
CC   as an antigen for detecting serum antibodies. The presence of serum
CC   antibodies to M. tuberculosis antigens in an individual indicates that
CC   the individual is infected with it. The fusion polypeptides are useful as
CC   sources of proteins for monitoring binding of serum antibodies to fusion
CC   proteins and as an immunogen to induce and/or enhance immune responses.
CC   The coding sequences can be ligated with a coding sequence of another
CC   molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC   can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC   represents the DNA encoding the Ra12 protein
XX
SQ   Sequence 396 BP; 68 A; 132 C; 134 G; 62 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                7.98e-55        Length:          396
Score:                    675.00          Matches:         132
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       4               Gaps:            0

US-09-684-215B-4 (1-132) x AAL40769 (1-396)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396

RESULT 2
AAL40772
ID   AAL40772 standard; DNA; 672 BP.
XX
AC   AAL40772;
XX
DT   06-AUG-2003  (revised)
DT   03-OCT-2002  (first entry)
XX
DE   Nucleotide sequence encoding Ra12-mammaglobin fusion protein.
XX
KW   Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW   vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW   immunogen; cytokine; gene; ds.
XX
OS   Mammalia.
OS   Unidentified.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .666
FT                   /*tag=  a
FT                   /product= "Ra12-mammaglobin fusion protein"
XX
PN   WO200125401-A2.
XX
PD   12-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US027652.
XX
PR   07-OCT-1999;   99US-0158585P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J;
XX
DR   WPI; 2001-266299/27.
DR   P-PSDB; AAO22141.
XX
PT   Recombinant nucleic acid molecule for producing high yield expression of
PT   desired fusion polypeptides, encodes fusion polypeptide comprising
PT   Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS   Disclosure; Fig 5; 39pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC   kDa C-terminal fragment of serine protease antigen MTB32A of
CC   Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC   The recombinant fusion nucleic acids and polypeptides are useful for
CC   providing stable and high yield expression of fusion polypeptides of both
CC   eukaryotic and prokaryotic origin and to encode a protein product for use
CC   as an antigen for detecting serum antibodies. The presence of serum
CC   antibodies to M. tuberculosis antigens in an individual indicates that
CC   the individual is infected with it. The fusion polypeptides are useful as
CC   sources of proteins for monitoring binding of serum antibodies to fusion
CC   proteins and as an immunogen to induce and/or enhance immune responses.
CC   The coding sequences can be ligated with a coding sequence of another
CC   molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC   can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC   represents the DNA encoding the Ra12-mammaglobin fusion protein. (Updated
CC   on 06-AUG-2003 to correct OS field.)
XX
SQ   Sequence 672 BP; 160 A; 185 C; 187 G; 140 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.45e-54        Length:          672
Score:                    675.00          Matches:         132
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       4               Gaps:            0

US-09-684-215B-4 (1-132) x AAL40772 (1-672)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84
```

```
Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 3
AAZ20206
ID AAZ20206 standard; DNA; 702 BP.
XX
AC AAZ20206;
XX
DT 17-JAN-2000 (first entry)
XX
DE Mycobacterium tuberculosis antigen fusion protein Mtb24 DNA.
XX
KW Tuberculosis; antigen; fusion protein; Mtb24; Ra12; DPPD; diagnosis;
KW therapy; vaccine; immunogen; ss.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9951748-A2.
XX
PD 14-OCT-1999.
XX
PF 07-APR-1999; 99WO-US007717.
XX
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Alderson M, Campos-Neto A;
XX
DR WPI; 1999-601610/51.
DR P-PSDB; AAY32071.
XX
PT New fusion proteins useful for diagnosis, prevention and treatment of
PT tuberculosis.
XX
PS Example; Fig 13A-B; 83pp; English.
XX
CC This DNA sequence includes a coding region for a recombinant
CC Mycobacterium tuberculosis bi-antigen fusion protein (see AAY32071),
CC termed Mtb24, composed of the antigens Ra12 and DPPD. The DNA is useful
CC for the recombinant production of the fusion protein. Coding sequences
CC for the antigens were modified by PCR in order to facilitate their fusion
CC and subsequent expression of the fusion protein, and then ligated. The
CC invention provides fusion proteins (see AAY32059-71) containing at least
CC 2 M. tuberculosis antigens. The new fusion proteins and polynucleotides
CC encoding them are useful as vaccines for preventing tuberculosis
CC (claimed), for diagnosis (via in vitro assays or intradermal skin tests
CC for detection of anti-M. tuberculosis antibodies), monitoring of disease
CC progression, and treatment of tuberculosis. They are more effective
CC immunogens than mixtures of the individual protein components
XX
SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.53e-54	Length:	702
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-684-215B-4 (1-132) x AAZ20206 (1-702)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 4
AAL40770
ID AAL40770 standard; DNA; 702 BP.
XX
AC AAL40770;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ra12-DPPD fusion protein.
XX
KW Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 4. .696
FT /*tag= a
FT /product= "Ra12-DPPD fusion protein"
XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027652.
XX
PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.

XX
PI Skeiky Y, Guderian J;
XX
DR WPI; 2001-266299/27.
DR P-PSDB; AAO22139.
XX
PT Recombinant nucleic acid molecule for producing high yield expression of
PT desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS Example 1; Fig 3; 39pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Ra12-DPPD fusion protein
XX
SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.53e-54	Length:	702
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-4 (1-132) x AAL40770 (1-702)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 5
ABK14140
ID ABK14140 standard; DNA; 702 BP.
XX
AC ABK14140;
XX
DT 29-AUG-2003 (revised)
DT 08-MAY-2002 (first entry)
XX
DE DNA encoding antigenic fusion protein Ra12-DPPD (Mtb24).
XX
KW Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;
KW tuberculostatic; immunogen; vaccine; Ra12-DPPD; Mtb24.
XX
OS Mycobacterium tuberculosis.
OS Chimeric.
XX

FH	Key	Location/Qualifiers
FT	CDS	1. .702
FT		/*tag= a
FT		/product= "Mtb24 #1"
FT		/partial
FT		/transl_except= (pos:694. .696, aa:Xaa)
FT		/note= "No start or stop codon. Xaa= In frame stop codon"
FT	CDS	2. .702
FT		/*tag= b
FT		/product= "Mtb24 #2"
FT		/partial
FT		/note= "No start or stop codon"
FT		/transl_except= (pos:263. .265, aa:Xaa) /transl_except=
FT		(pos:353. .355, aa:Xaa) /transl_except= (pos:395. .397,
FT		aa:Xaa) /transl_except= (pos:470. .472, aa:Xaa)
FT		/transl_except= (pos:701. .702, aa:Ser)
FT		/note= "This codon has an apparent 1 nucleotide deletion
FT		which alters the reading frame. Xaa= In frame stop codon"
FT	CDS	3. .701
FT		/*tag= c
FT		/product= "Mtb24 #3"
FT		/partial
FT		/note= "No start or stop codon"
FT		/transl_except= (pos:1. .2, aa:Pro) /transl_except=
FT		(pos:39. .41, aa:Xaa) /transl_except= (pos:321. .323,
FT		aa:Xaa) /transl_except= (pos:339. .341, aa:Xaa)
FT		/transl_except= (pos:450. .452, aa:Xaa) /transl_except=
FT		(pos:621. .623, aa:Xaa)
FT		/note= "No start or stop codon. Xaa= In frame stop codon"

XX
PN US2002009459-A1.
XX
PD 24-JAN-2002.
XX
PF 07-APR-1999; 99US-00287849.
XX
PR 13-MAR-1997; 97US-00818112.
PR 01-OCT-1997; 97US-00942578.
PR 18-FEB-1998; 98US-00025197.
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (REED/) REED S G.
PA (SKEI/) SKEIKY Y A.
PA (DILL/) DILLON D C.
PA (ALDE/) ALDERSON M.
PA (CAMP/) CAMPOS-NETO A.
XX
PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;
XX
DR WPI; 2002-171134/22.
DR P-PSDB; AAU74600, AAU76541, AAU76542.
XX
PT New fusion proteins of Mycobacterium tuberculosis antigens, useful for
PT diagnosing, treating or preventing M. tuberculosis infection,
PT particularly as vaccine for treating or preventing tuberculosis.
XX
PS Example; Fig 13; 62pp; English.
XX
CC The invention relates to a purified polypeptide which induces an immune

CC response of Mycobacterium tuberculosis. Polypeptides of the invention are
CC useful for diagnosing, treating or preventing M. tuberculosis infection,
CC particularly tuberculosis infection. In particular, the polypeptides are
CC useful as a vaccine formulation with an adjuvant to afford long-term
CC protection in animals against the development of tuberculosis. The
CC protein coding sequence may be used to encode a protein product for use
CC as an immunogen to induce and/or enhance an immune response to M.
CC tuberculosis. This sequence represents DNA encoding an M. tuberculosis
CC fusion protein of the invention. This polynucleotide encodes 3 different
CC proteins, each in a different reading frame. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.53e-54	Length:	702
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-4 (1-132) x ABK14140 (1-702)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 6
AAD47078
ID AAD47078 standard; DNA; 1002 BP.
XX
AC AAD47078;
XX
DT 27-JAN-2003 (first entry)
XX
DE Mycobacterium tuberculosis Ra35FLMutSA mutant antigenic protein DNA.
XX
KW Vaccine; immunity; diagnostic agent; gene therapy; Ra35FLMutSA; mutant;
KW gene; antigen; ds.
XX
OS Mycobacterium tuberculosis.
OS Synthetic.
XX

FH	Key	Location/Qualifiers
FT	CDS	4. .996
FT		/*tag= a
FT		/product= "Ra35FLMutSA mutant antigenic protein"

XX
PN WO200272792-A2.
XX
PD 19-SEP-2002.
XX
PF 13-MAR-2002; 2002WO-US008223.
XX
PR 13-MAR-2001; 2001US-0275837P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Brannon M, Guderian J;
XX
DR WPI; 2002-759844/82.
DR P-PSDB; AAE29703.
XX
PT New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT tuberculosis.
XX
PS Disclosure; Page 80-81; 155pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC polynucleotide sequence encoding an antigen or an antigenic fragment from
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC polypeptide or its fragment. The Leishmania polynucleotide is selected
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides, as
CC in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is M.
CC tuberculosis Ra35FLMutSA mutant antigenic protein encoding DNA
XX
SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.29e-54	Length:	1002
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-4 (1-132) x AAD47078 (1-1002)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
```

```
                    ||||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993

RESULT 7
AAD47077
ID   AAD47077 standard; DNA; 1002 BP.
XX
AC   AAD47077;
XX
DT   27-JAN-2003  (first entry)
XX
DE   Mycobacterium tuberculosis mature Ra35 antigen encoding DNA.
XX
KW   Vaccine; immunity; diagnostic agent; gene therapy; Ra35 antigen; gene;
KW   ds.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .996
FT                   /*tag= a
FT                   /product= "Ra35 mature antigenic protein"
FT                   /transl_except= (pos:547. .549, aa:Asp)
FT                   /transl_except= (pos:550. .552, aa:Ser)
XX
PN   WO200272792-A2.
XX
PD   19-SEP-2002.
XX
PF   13-MAR-2002; 2002WO-US008223.
XX
PR   13-MAR-2001; 2001US-0275837P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Brannon M,  Guderian J;
XX
DR   WPI; 2002-759844/82.
DR   P-PSDB; AAE29702.
XX
PT   New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT   M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT   against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT   tuberculosis.
XX
PS   Disclosure; Page 79; 155pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC   polynucleotide sequence encoding an antigen or an antigenic fragment from
CC   Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC   polypeptide or its fragment. The Leishmania polynucleotide is selected
CC   from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC   are used in methods for eliciting immune response in mammals. They are
CC   useful as vaccines to elicit protective immunity against pathogenic
CC   microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC   polypeptides are used for enhancing the expression of polynucleotides, as
CC   in vivo diagnostic agents and for raising antibodies in a non-human
CC   animal. The invention is used in gene therapy. The present sequence is M.
CC   tuberculosis mature Ra35 (N-terminus of MTB32A; RA35FL) antigen encoding
CC   DNA
XX
SQ   Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.29e-54        Length:         1002
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-4 (1-132) x AAD47077 (1-1002)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993

RESULT 8
AAD28336
ID   AAD28336 standard; cDNA; 1002 BP.
XX
AC   AAD28336;
XX
DT   22-APR-2002  (first entry)
XX
DE   Mycobacterium sp. MTB32A (Ra35FL) mature protein encoding cDNA.
XX
KW   Fusion protein; antigen; serological sensitivity; immune response;
KW   tuberculosis; infection; vaccine; MTB32A; Ra32FL; ss.
XX
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .996
FT                   /*tag= a
FT                   /product= "Ra35 mature protein"
FT                   /transl_except= (pos:547. .549, aa:Asp)
FT                   /transl_except= (pos:550. .552, aa:Ser)
XX
PN   WO200198460-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019959.
XX
PR   20-JUN-2000; 2000US-00597796.
PR   01-FEB-2001; 2001US-0265737P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Reed S,  Alderson M;
XX
DR   WPI; 2002-147798/19.
DR   P-PSDB; AAE17566.
XX
PT   Composition comprising MTB39 antigen and MTB32A antigen from
PT   Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS   Disclosure; Page 95; 136pp; English.
XX
CC   The present invention relates to fusion proteins containing at least two
```

```
CC   Mycobacterium species antigens, nucleotides encoding them and
CC   compositions comprising such fusion proteins. The present invention
CC   particularly relates to nucleic acids encoding fusion proteins that
CC   include two or more individual M. tuberculosis antigens which increase
CC   the serological sensitivity of sera from individuals infected with
CC   tuberculosis and methods for their use in diagnosis, prevention and
CC   treatment of tuberculosis infection. Sequences of the invention are
CC   useful for eliciting an immune response in a mammal, e.g., human,
CC   immunised with BCG. They are useful in the diagnosis, treatment and
CC   prevention of Mycobacterium infection. The fusion proteins and the
CC   polynucleotides are useful as diagnostic tools in patients infected with
CC   Mycobacterium, in vitro and in vivo assays for detecting humoral
CC   antibodies or cell-mediated immunity against M. tuberculosis, for the
CC   diagnosis of an infection or monitoring of disease progression, as
CC   immunogens to generate or elicit a protective immune response in a
CC   patient and for raising anti-M. tuberculosis antibodies in a non-human
CC   animal. Sequences of the invention are also used as vaccines. MTB32A
CC   fusion proteins of the invention are useful as in vivo diagnostic agents
CC   for intradermal skin test. The present sequence is a cDNA encoding
CC   Mycobacterium species MTB32A (Ra32FL) mature protein
XX
SQ   Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.29e-54        Length:         1002
Score:                    675.00          Matches:        132
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-4 (1-132) x AAD28336 (1-1002)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993

RESULT 9
AAD28337
ID   AAD28337 standard; cDNA; 1002 BP.
XX
AC   AAD28337;
XX
DT   22-APR-2002  (first entry)
XX
DE   Mycobacterium species Ra35FLMutSA mutant cDNA.
XX
KW   Fusion protein; antigen; serological sensitivity; immune response;
KW   tuberculosis; infection; vaccine; Ra32FLMutSA; mutant; mutein; ss.
XX
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .996
FT                   /*tag=  a
FT                   /product= "Ra35FLMutSA protein"
XX
PN   WO200198460-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019959.
XX
PR   20-JUN-2000; 2000US-00597796.
PR   01-FEB-2001; 2001US-0265737P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Reed S,  Alderson M;
XX
DR   WPI; 2002-147798/19.
DR   P-PSDB; AAB17567.
XX
PT   Composition comprising MTB39 antigen and MTB32A antigen from
PT   Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS   Claim 67; Page 96-97; 136pp; English.
XX
CC   The present invention relates to fusion proteins containing at least two
CC   Mycobacterium species antigens, nucleotides encoding them and
CC   compositions comprising such fusion proteins. The present invention
CC   particularly relates to nucleic acids encoding fusion proteins that
CC   include two or more individual M. tuberculosis antigens which increase
CC   the serological sensitivity of sera from individuals infected with
CC   tuberculosis and methods for their use in diagnosis, prevention and
CC   treatment of tuberculosis infection. Sequences of the invention are
CC   useful for eliciting an immune response in a mammal, e.g., human,
CC   immunised with BCG. They are useful in the diagnosis, treatment and
CC   prevention of Mycobacterium infection. The fusion proteins and the
CC   polynucleotides are useful as diagnostic tools in patients infected with
CC   Mycobacterium, in vitro and in vivo assays for detecting humoral
CC   antibodies or cell-mediated immunity against M. tuberculosis, for the
CC   diagnosis of an infection or monitoring of disease progression, as
CC   immunogens to generate or elicit a protective immune response in a
CC   patient and for raising anti-M. tuberculosis antibodies in a non-human
CC   animal. Sequences of the invention are also used as vaccines. MTB32A
CC   fusion proteins of the invention are useful as in vivo diagnostic agents
CC   for intradermal skin test. The present sequence is a cDNA encoding
CC   Mycobacterium species Ra32FL mature protein mutant, Ra32FLMutSA
XX
SQ   Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.29e-54        Length:         1002
Score:                    675.00          Matches:        132
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-4 (1-132) x AAD28337 (1-1002)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993

RESULT 10
AAX34251
ID   AAX34251 standard; DNA; 1068 BP.
XX
AC   AAX34251;
XX
DT   06-JUL-1999  (first entry)
XX
DE   Mycobacterium species nucleic acid sequence 50D.
XX
KW   Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW   hybridisation; detection; vaccine; immunisation; infection; ss.
XX
OS   Mycobacterium sp.
XX
PN   WO9909186-A2.
XX
PD   25-FEB-1999.
XX
PF   14-AUG-1998;   98WO-FR001813.
XX
PR   14-AUG-1997;   97FR-00010404.
PR   11-SEP-1997;   97FR-00011325.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Gicquel B,  Portnoie D,  Lim E,  Pelicic V,  Guigueno A;
PI   Goguet De La Salmoniere Y;
XX
DR   WPI; 1999-181045/15.
DR   P-PSDB; AAY05000.
XX
PT   Mycobacterial DNA vectors containing reporter constructs - for
PT   identifying coding or promoter sequences involved in infection-associated
PT   protein expression.
XX
PS   Claim 22; Fig 50D; 309pp; French.
XX
CC   Sequences AAX34001-X34252 represent nucleic acids encoding secreted
CC   proteins from various Mycobacterium species microorganisms. The
CC   nucleotide sequences can be used as primers and probes for methods for
CC   detecting and identifying mycobacteria, especially belonging to the M.
CC   tuberculosis complex. The encoded proteins can be used in vaccines for
CC   immunisation against a bacterial or viral infection
XX
SQ   Sequence 1068 BP; 169 A; 352 C; 375 G; 172 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.46e-54        Length:         1068
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     2               Gaps:           0

US-09-684-215B-4 (1-132) x AAX34251 (1-1068)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 729

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 789

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 849

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 909

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 969

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1029

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db       1030 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1065

RESULT 11
AAX34252
ID   AAX34252 standard; DNA; 1143 BP.
XX
AC   AAX34252;
XX
DT   06-JUL-1999  (first entry)
XX
DE   Mycobacterium species nucleic acid sequence 50F.
XX
KW   Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW   hybridisation; detection; vaccine; immunisation; infection; ss.
XX
OS   Mycobacterium sp.
XX
PN   WO9909186-A2.
XX
PD   25-FEB-1999.
XX
PF   14-AUG-1998;   98WO-FR001813.
XX
PR   14-AUG-1997;   97FR-00010404.
PR   11-SEP-1997;   97FR-00011325.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Gicquel B,  Portnoie D,  Lim E,  Pelicic V,  Guigueno A;
PI   Goguet De La Salmoniere Y;
XX
DR   WPI; 1999-181045/15.
DR   P-PSDB; AAY04830.
XX
PT   Mycobacterial DNA vectors containing reporter constructs - for
PT   identifying coding or promoter sequences involved in infection-associated
PT   protein expression.
XX
PS   Claim 22; Fig 50F; 309pp; French.
XX
CC   Sequences AAX34001-X34252 represent nucleic acids encoding secreted
CC   proteins from various Mycobacterium species microorganisms. The
CC   nucleotide sequences can be used as primers and probes for methods for
CC   detecting and identifying mycobacteria, especially belonging to the M.
CC   tuberculosis complex. The encoded proteins can be used in vaccines for
```

```
CC   immunisation against a bacterial or viral infection
XX
SQ   Sequence 1143 BP; 189 A; 373 C; 395 G; 186 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.66e-54        Length:         1143
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     2               Gaps:           0

US-09-684-215B-4 (1-132) x AAX34252 (1-1143)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 804

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 864

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 924

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 984

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1044

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1104

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db       1105 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1140

RESULT 12
AAS03793
ID   AAS03793 standard; DNA; 1629 BP.
XX
AC   AAS03793;
XX
DT   29-AUG-2001  (first entry)
XX
DE   M. tuberculosis DNA encoding antigen HTCC#1 fusion protein #4.
XX
KW   TbRA12-HTCC#1; antigen; vaccine; tuberculosis; AIDS;
KW   acquired immunodeficiency disease; His Tag; ds.
XX
OS   Mycobacterium tuberculosis.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1629
FT                   /*tag=  a
FT                   /product= "TbRA12-HTCC#1"
FT                   /transl_except= (pos:1621. .1623,aa:Xaa)
FT                   /note= "Xaa= In frame STOP codon"
FT                   /partial
FT                   /note= "No start or stop codon. Although the sequence
FT                   does contain an in frame stop codon, 2 further amino
FT                   acids are shown in Figure 8 as being encoded by the
FT                   present sequence, without a further stop codon"
FT   misc_feature    25. .426
FT                   /*tag=  b
FT                   /note= "Region derived from TbRA12"
FT   misc_feature    427. .444
FT                   /*tag=  c
FT                   /note= "Region derived from Thrombin"
FT   misc_feature    445. .1629
FT                   /*tag=  d
FT                   /note= "Region derived from HTCC#1"
FT
XX
PN   WO200124820-A1.
XX
PD   12-APR-2001.
XX
PF   10-OCT-2000; 2000WO-US028095.
XX
PR   07-OCT-1999;   99US-0158338P.
PR   07-OCT-1999;   99US-0158425P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Reed S,  Houghton RL,  Mcneill PD,  Dillon DC,  Lodes ML;
XX
DR   WPI; 2001-290576/30.
XX
PT   Vaccinating against Mycobacteria infections in mammals using fusion
PT   proteins comprising combinations of heterologous antigens.
XX
PS   Example 2; Fig 8; 168pp; English.
XX
CC   The sequence encodes Mycobacterium tuberculosis fusion protein, TbRA12-
CC   HTCC#1 and includes a His tag at the N-terminus to aid purification.
CC   Compositions comprising at least 2 heterologous antigens, as a fusion
CC   protein, and vectors expressing the fusion proteins are used as vaccines
CC   to prophylactically immunise mammals (especially humans) against
CC   infection by Mycobacteria. The compositions contain at least 2
CC   heterologous antigens that increase the serological sensitivity of
CC   individuals infected with tuberculosis, a disease frequently affecting
CC   patients with acquired immunodeficiency disease, AIDS
XX
SQ   Sequence 1629 BP; 298 A; 500 C; 535 G; 296 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.97e-54        Length:         1629
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     5               Gaps:           0

US-09-684-215B-4 (1-132) x AAS03793 (1-1629)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCCGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
```

```
                    ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 13
AAL40771
ID   AAL40771 standard; DNA; 1742 BP.
XX
AC   AAL40771;
XX
DT   03-OCT-2002  (first entry)
XX
DE   Nucleotide sequence encoding Ra12-WT1 fusion protein.
XX
KW   Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW   vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW   immunogen; cytokine; gene; ds.
XX
OS   Unidentified.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .1740
FT                   /*tag=  a
FT                   /product= "Ra12-WT1 fusion protein"
XX
PN   WO200125401-A2.
XX
PD   12-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US027652.
XX
PR   07-OCT-1999;   99US-0158585P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J;
XX
DR   WPI; 2001-266299/27.
DR   P-PSDB; AAO22140.
XX
PT   Recombinant nucleic acid molecule for producing high yield expression of
PT   desired fusion polypeptides, encodes fusion polypeptide comprising
PT   Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS   Disclosure; Fig 4; 39pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC   kDa C-terminal fragment of serine protease antigen MTB32A of
CC   Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC   The recombinant fusion nucleic acids and polypeptides are useful for
CC   providing stable and high yield expression of fusion polypeptides of both
CC   eukaryotic and prokaryotic origin and to encode a protein product for use
CC   as an antigen for detecting serum antibodies. The presence of serum
CC   antibodies to M. tuberculosis antigens in an individual indicates that
CC   the individual is infected with it. The fusion polypeptides are useful as
CC   sources of proteins for monitoring binding of serum antibodies to fusion
CC   proteins and as an immunogen to induce and/or enhance immune responses.
CC   The coding sequences can be ligated with a coding sequence of another
CC   molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC   can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC   represents the DNA encoding the Ra12-WT1 fusion protein
XX
SQ   Sequence 1742 BP; 372 A; 569 C; 490 G; 311 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.28e-54        Length:          1742
Score:                    675.00          Matches:         132
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       4               Gaps:            0

US-09-684-215B-4 (1-132) x AAL40771 (1-1742)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 14
AAL40768
ID   AAL40768 standard; DNA; 1871 BP.
XX
AC   AAL40768;
XX
DT   03-OCT-2002  (first entry)
XX
DE   Nucleotide sequence encoding MTB32A protein.
XX
KW   Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW   vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW   immunogen; cytokine; gene; ds.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   CDS             89. .1156
FT                   /*tag=  a
FT                   /product= "Bacillus mycobacterium MTB32A protein"
XX
PN   WO200125401-A2.
XX
PD   12-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US027652.
XX
PR   07-OCT-1999;   99US-0158585P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J;
XX
DR   WPI; 2001-266299/27.
DR   P-PSDB; AAO22137.
XX
PT   Recombinant nucleic acid molecule for producing high yield expression of
PT   desired fusion polypeptides, encodes fusion polypeptide comprising
PT   Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS   Disclosure; Fig 1; 39pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
```

CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Mycobacterium tuberculosis MTB32A protein
XX
SQ Sequence 1871 BP; 317 A; 617 C; 605 G; 332 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.65e-54	Length:	1871
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-4 (1-132) x AAL40768 (1-1871)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153
```

RESULT 15
ADA26353
ID ADA26353 standard; DNA; 2181 BP.
XX
AC ADA26353;
XX
DT 20-NOV-2003 (first entry)
XX
DE Mycobacterium MTB32-MTB39F fusion protein encoding DNA.
XX
KW ds; gene; fusion polypeptide; MTB32A; MTB39; antigen; MTB32A; MTB39;
KW MTB85A; tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS Chimeric.
OS Mycobacterium sp.
XX
FH Key Location/Qualifiers
FT CDS 4. .2175
FT /*tag= a
FT /product= "MTB32-MTB39F fusion polypeptide"
XX
PN WO2003070187-A2.
XX
PD 28-AUG-2003.
XX
PF 18-FEB-2003; 2003WO-US004903.
XX
PR 15-FEB-2002; 2002US-0357351P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J, Reed S;
XX
DR WPI; 2003-697554/66.
DR P-PSDB; ADA26354.
XX
PT New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT and MTB39 antigens, with or without the MTB85A antigen, from a
PT Mycobacterium species, useful for diagnosing, preventing and/or treating
PT tuberculosis infection.
XX
PS Claim 1; Fig 1; 112pp; English.
XX
CC The invention relates to a novel nucleic acid encoding a fusion
CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC A polypeptide of the invention has tuberculostatic activity. A
CC polynucelotide of the invention may have a use in gene therapy, and as a
CC vaccine. The methods and compositions of the present invention are useful
CC for diagnosing, preventing and/or treating tuberculosis infection. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 2181 BP; 341 A; 693 C; 792 G; 355 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.53e-54	Length:	2181
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0

US-09-684-215B-4 (1-132) x ADA26353 (1-2181)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
```

```
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993

RESULT 16
AAD47084
ID   AAD47084 standard; DNA; 2190 BP.
XX
AC   AAD47084;
XX
DT   29-AUG-2003  (revised)
DT   27-JAN-2003  (first entry)
XX
DE   Mycobacterium sp. MTB72FMutSA fusion protein encoding DNA.
XX
KW   Vaccine; immunity; diagnostic agent; gene therapy; TbH9; antigen;
KW   Ra35MutSA; Ra12; MTB72MutSA; chimeric; gene; ds.
XX
OS   Mycobacterium sp.
OS   Mycobacterium tuberculosis.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2190
FT                   /*tag= a
FT                   /product= "MTB72FMutSA fusion protein"
XX
PN   WO200272792-A2.
XX
PD   19-SEP-2002.
XX
PF   13-MAR-2002; 2002WO-US008223.
XX
PR   13-MAR-2001; 2001US-0275837P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Brannon M,  Guderian J;
XX
DR   WPI; 2002-759844/82.
DR   P-PSDB; AAE29709.
XX
PT   New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT   M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT   against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT   tuberculosis.
XX
PS   Disclosure; Page 92-93; 155pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC   polynucleotide sequence encoding an antigen or an antigenic fragment from
CC   Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC   polypeptide or its fragment. The Leishmania polynucleotide is selected
CC   from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC   are used in methods for eliciting immune response in mammals. They are
CC   useful as vaccines to elicit protective immunity against pathogenic
CC   microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC   polypeptides are used for enhancing the expression of polynucleotides, as
CC   in vivo diagnostic agents and for raising antibodies in a non-human
CC   animal. The invention is used in gene therapy. The present sequence is a
CC   DNA encoding MTB72F fusion protein. This fusion protein comprises
CC   Ra35MutSA mutant protein and Ra12 protein from M. tuberculosis and TbH9
CC   protein from Mycobacterium sp. (Updated on 29-AUG-2003 to standardise OS
CC   field)
XX
SQ   Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.55e-54      Length:          2190
Score:                  675.00        Matches:         132
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     6             Gaps:            0

US-09-684-215B-4 (1-132) x AAD47084 (1-2190)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 141

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 201

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 321

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 381

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        382 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 417

RESULT 17
AAD28343
ID   AAD28343 standard; DNA; 2190 BP.
XX
AC   AAD28343;
XX
DT   22-APR-2002  (first entry)
XX
DE   Mycobacterium species MTB72FMutSA fusion protein encoding DNA.
XX
KW   Fusion protein; antigen; serological sensitivity; immune response;
KW   tuberculosis; infection; vaccine; MTB72FMutSA; Ra12-TbH9-Ra35MutSA;
KW   mutant; mutein; ds.
XX
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2190
FT                   /*tag= a
FT                   /product= "MTB72FMutSA fusion protein"
FT   misc_feature    22. .417
FT                   /*tag= b
FT                   /note= "Ra12 DNA fragment"
FT   misc_feature    424. .1596
FT                   /*tag= c
FT                   /note= "TbH9FL DNA fragment"
FT   misc_feature    1603. .2187
FT                   /*tag= d
FT                   /note= "Ra35 DNA fragment"
FT   mutation        replace(2128, T)
FT                   /*tag= e
XX
PN   WO200198460-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019959.
XX
PR   20-JUN-2000; 2000US-00597796.
PR   01-FEB-2001; 2001US-0265737P.
XX
PA   (CORI-) CORIXA CORP.
```

XX
PI Skeiky Y, Reed S, Alderson M;
XX
DR WPI; 2002-147798/19.
DR P-PSDB; AAE17573.
XX
PT Composition comprising MTB39 antigen and MTB32A antigen from
PT Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS Claim 81; Page 108-109; 136pp; English.
XX
CC The present invention relates to fusion proteins containing at least two
CC Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected with
CC Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is a DNA encoding
CC Mycobacterium species MTB72FMutSA (Ra12-TbH9-Ra35MutSA) mutant fusion
CC protein
XX
SQ Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.55e-54	Length:	2190
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-4 (1-132) x AAD28343 (1-2190)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 141

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 201

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 321

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 381

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db        382 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 417
```

RESULT 18
AAL40773
ID AAL40773 standard; DNA; 2191 BP.
XX
AC AAL40773;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ra12-H9-32A fusion protein.
XX
KW Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX

FH	Key	Location/Qualifiers
FT	CDS	1. .2190
FT		/*tag= a
FT		/product= "Ra12-H9-32A fusion protein"

XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027652.
XX
PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J;
XX
DR WPI; 2001-266299/27.
DR P-PSDB; AAO22142.
XX
PT Recombinant nucleic acid molecule for producing high yield expression of
PT desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS Disclosure; Fig 6; 39pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Ra12-H9-32A fusion protein
XX
SQ Sequence 2191 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	5.56e-54	Length:	2191
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-4 (1-132) x AAL40773 (1-2191)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 141

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 201

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 321

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 381

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||
Db        382 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 417
```

RESULT 19
ABK14128
ID ABK14128 standard; DNA; 2286 BP.
XX
AC ABK14128;
XX
DT 29-AUG-2003 (revised)
DT 08-MAY-2002 (first entry)
XX
DE DNA encoding antigenic fusion protein Ra12-TbH9-Ra35 (Mtb32-Mtb39).
XX
KW Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;
KW tuberculostatic; immunogen; vaccine; Mtb32-Mtb39; Ra12; TbH9; Ra35;
KW Ra12-TbH9-Ra35.
XX
OS Mycobacterium tuberculosis.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 42. .2231
FT /*tag= a
FT /product= "Mtb32-Mtb39"
FT /transl_except= (pos:498. .506, aa:Asn-Ala)
FT /transl_except= (pos:597. .605, aa:Ala-Gln)
FT /transl_except= (pos:798. .802, aa:Ala)
FT /note= "This codon has an apparent 2 nucleotide insertion
FT which alters the reading frame"
XX
PN US2002009459-A1.
XX
PD 24-JAN-2002.
XX
PF 07-APR-1999; 99US-00287849.
XX
PR 13-MAR-1997; 97US-00818112.
PR 01-OCT-1997; 97US-00942578.
PR 18-FEB-1998; 98US-00025197.
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (REED/) REED S G.
PA (SKEI/) SKEIKY Y A.
PA (DILL/) DILLON D C.
PA (ALDE/) ALDERSON M.
PA (CAMP/) CAMPOS-NETO A.
XX
PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;
XX
DR WPI; 2002-171134/22.
DR P-PSDB; AAU74588.
XX
PT New fusion proteins of Mycobacterium tuberculosis antigens, useful for
PT diagnosing, treating or preventing M. tuberculosis infection,
PT particularly as vaccine for treating or preventing tuberculosis.
XX
PS Example; Fig 1; 62pp; English.
XX
CC The invention relates to a purified polypeptide which induces an immune
CC response of Mycobacterium tuberculosis. Polypeptides of the invention are
CC useful for diagnosing, treating or preventing M. tuberculosis infection,
CC particularly tuberculosis infection. In particular, the polypeptides are
CC useful as a vaccine formulation with an adjuvant to afford long-term
CC protection in animals against the development of tuberculosis. The
CC protein coding sequence may be used to encode a protein product for use
CC as an immunogen to induce and/or enhance an immune response to M.
CC tuberculosis. This sequence represents DNA encoding an M. tuberculosis
CC fusion protein of the invention. (Updated on 29-AUG-2003 to standardise
CC OS field)
XX
SQ Sequence 2286 BP; 373 A; 719 C; 807 G; 381 T; 0 U; 6 Other;

Alignment Scores:

Pred. No.:	5.83e-54	Length:	2286
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-4 (1-132) x ABK14128 (1-2286)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

RESULT 20
AAZ20194
ID AAZ20194 standard; DNA; 2287 BP.
XX
AC AAZ20194;
XX
DT 17-JAN-2000 (first entry)
XX
DE Mycobacterium tuberculosis antigen fusion protein Mtb32A DNA.

```
XX
KW   Tuberculosis; antigen; fusion protein; Mtb32A; Ra12; TbH9; Ra35;
KW   diagnosis; therapy; vaccine; immunogen; ss.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   CDS             42. .2231
FT                   /*tag=  a
XX
PN   WO9951748-A2.
XX
PD   14-OCT-1999.
XX
PF   07-APR-1999;    99WO-US007717.
XX
PR   07-APR-1998;    98US-00056556.
PR   30-DEC-1998;    98US-00223040.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Alderson M,  Campos-Neto A;
XX
DR   WPI; 1999-601610/51.
DR   P-PSDB; AAY32059.
XX
PT   New fusion proteins useful for diagnosis, prevention and treatment of
PT   tuberculosis.
XX
PS   Example; Fig 1A-B; 83pp; English.
XX
CC   This DNA sequence includes a coding region for a recombinant
CC   Mycobacterium tuberculosis tri-antigen fusion protein (see AAY32059),
CC   termed Mtb32A, composed of the antigens Ra12, TbH9 and Ra35. The DNA is
CC   useful for the recombinant production of the fusion protein. Coding
CC   sequences for the antigens were modified by PCR in order to facilitate
CC   their fusion and subsequent expression of the fusion protein. 3 Coding
CC   sequences for Ra12, TbH9 and Ra25 were ligated to encode Mtb32A. The
CC   invention provides fusion proteins (see AAY32059-71) containing at least
CC   2 M. tuberculosis antigens. The new fusion proteins and polynucleotides
CC   encoding them are useful as vaccines for preventing tuberculosis
CC   (claimed), for diagnosis (via in vitro assays or intradermal skin tests
CC   for detection of anti-M. tuberculosis antibodies), monitoring of disease
CC   progression, and treatment of tuberculosis. They are more effective
CC   immunogens than mixtures of the individual protein components
XX
SQ   Sequence 2287 BP; 374 A; 723 C; 805 G; 382 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:              5.83e-54        Length:         2287
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     2               Gaps:           0

US-09-684-215B-4 (1-132) x AAZ20194 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458

RESULT 21
AAD47083
ID   AAD47083 standard; DNA; 2287 BP.
XX
AC   AAD47083;
XX
DT   29-AUG-2003  (revised)
DT   27-JAN-2003  (first entry)
XX
DE   Mycobacterium sp. MTB72F fusion protein encoding DNA.
XX
KW   Vaccine; immunity; diagnostic agent; gene therapy; TbH9; antigen; Ra35;
KW   Ra12; MTB72F; chimeric; gene; ds.
XX
OS   Mycobacterium sp.
OS   Mycobacterium tuberculosis.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             42. .2231
FT                   /*tag=  a
FT                   /product= "MTB72F fusion protein"
XX
PN   WO200272792-A2.
XX
PD   19-SEP-2002.
XX
PF   13-MAR-2002; 2002WO-US008223.
XX
PR   13-MAR-2001; 2001US-0275837P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Brannon M,  Guderian J;
XX
DR   WPI; 2002-759844/82.
DR   P-PSDB; AAE29708.
XX
PT   New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT   M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT   against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT   tuberculosis.
XX
PS   Disclosure; Page 87-90; 155pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC   polynucleotide sequence encoding an antigen or an antigenic fragment from
CC   Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC   polypeptide or its fragment. The Leishmania polynucleotide is selected
CC   from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC   are used in methods for eliciting immune response in mammals. They are
CC   useful as vaccines to elicit protective immunity against pathogenic
CC   microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC   polypeptides are used for enhancing the expression of polynucleotides, as
CC   in vivo diagnostic agents and for raising antibodies in a non-human
CC   animal. The invention is used in gene therapy. The present sequence is a
CC   DNA encoding MTB72F fusion protein. This fusion protein comprises Ra12
CC   and Ra35 protein from Mycobacterium tuberculosis and TbH9 protein from
CC   Mycobacterium sp. (Updated on 29-AUG-2003 to standardise OS field)
```

```
XX
SQ   Sequence 2287 BP; 372 A; 717 C; 814 G; 381 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:              5.83e-54        Length:         2287
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-4 (1-132) x AAD47083 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458

RESULT 22
AAD28342
ID   AAD28342 standard; DNA; 2287 BP.
XX
AC   AAD28342;
XX
DT   22-APR-2002  (first entry)
XX
DE   Mycobacterium species MTB72F fusion protein encoding DNA.
XX
KW   Fusion protein; antigen; serological sensitivity; immune response;
KW   tuberculosis; infection; vaccine; MTB72F; Ra12-TbH9-Ra35; ds.
XX
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             42. .2231
FT                   /*tag=  a
FT                   /product= "MTB72F fusion protein"
FT   misc_feature    63. .458
FT                   /*tag=  b
FT                   /note= "Ra12 DNA fragment"
FT   misc_feature    465. .1637
FT                   /*tag=  c
FT                   /note= "TbH9FL DNA fragment"
FT   misc_feature    1644. .2228
FT                   /*tag=  d
FT                   /note= "Ra35 DNA fragment"
XX
PN   WO200198460-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019959.
XX
PR   20-JUN-2000; 2000US-00597796.
PR   01-FEB-2001; 2001US-0265737P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Reed S,  Alderson M;
XX
DR   WPI; 2002-147798/19.
DR   P-PSDB; AAE17572.
XX
PT   Composition comprising MTB39 antigen and MTB32A antigen from
PT   Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS   Claim 62; Page 103-106; 136pp; English.
XX
CC   The present invention relates to fusion proteins containing at least two
CC   Mycobacterium species antigens, nucleotides encoding them and
CC   compositions comprising such fusion proteins. The present invention
CC   particularly relates to nucleic acids encoding fusion proteins that
CC   include two or more individual M. tuberculosis antigens which increase
CC   the serological sensitivity of sera from individuals infected with
CC   tuberculosis and methods for their use in diagnosis, prevention and
CC   treatment of tuberculosis infection. Sequences of the invention are
CC   useful for eliciting an immune response in a mammal, e.g., human,
CC   immunised with BCG. They are useful in the diagnosis, treatment and
CC   prevention of Mycobacterium infection. The fusion proteins and the
CC   polynucleotides are useful as diagnostic tools in patients infected with
CC   Mycobacterium, in vitro and in vivo assays for detecting humoral
CC   antibodies or cell-mediated immunity against M. tuberculosis, for the
CC   diagnosis of an infection or monitoring of disease progression, as
CC   immunogens to generate or elicit a protective immune response in a
CC   patient and for raising anti-M. tuberculosis antibodies in a non-human
CC   animal. Sequences of the invention are also used as vaccines. MTB32A
CC   fusion proteins of the invention are useful as in vivo diagnostic agents
CC   for intradermal skin test. The present sequence is a DNA encoding
CC   Mycobacterium species MTB72F (Ra12-TbH9-Ra35) fusion protein
XX
SQ   Sequence 2287 BP; 372 A; 717 C; 814 G; 381 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:              5.83e-54        Length:         2287
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-4 (1-132) x AAD28342 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
                  ||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458

RESULT 23
ADA26360
ID   ADA26360 standard; DNA; 2451 BP.
XX
AC   ADA26360;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB72F-DPV (fusion MTB81F) protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2445
FT                   /*tag=  a
FT                   /product= " MTB72F-DPV (fusion MTB81F) protein"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26367.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 84; Fig 8; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 2451 BP; 389 A; 785 C; 870 G; 407 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.31e-54        Length:          2451
Score:                    675.00          Matches:         132
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       8               Gaps:            0

US-09-684-215B-4 (1-132) x ADA26360 (1-2451)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 24
ADA26359
ID   ADA26359 standard; DNA; 2487 BP.
XX
AC   ADA26359;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB72F-MTI (fusion MTB83F) protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2481
FT                   /*tag=  a
FT                   /product= "MTB72F-MTI (fusion MTB83F) protein"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26366.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 84; Fig 7; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
```

```
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 2487 BP; 394 A; 784 C; 896 G; 413 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.42e-54        Length:         2487
Score:                    675.00          Matches:        132
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       8               Gaps:           0

US-09-684-215B-4 (1-132) x ADA26359 (1-2487)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 25
ADA26358
ID   ADA26358 standard; DNA; 2637 BP.
XX
AC   ADA26358;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB72F-Erd14 (fusion MTB89F) protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2631
FT                   /*tag=  a
FT                   /product= "MTB72F-Erd14 (fusion MTB89F)"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26365.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 84; Fig 6; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 2637 BP; 428 A; 840 C; 928 G; 441 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.86e-54        Length:         2637
Score:                    675.00          Matches:        132
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       8               Gaps:           0

US-09-684-215B-4 (1-132) x ADA26358 (1-2637)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 26
AAD47110
ID   AAD47110 standard; DNA; 2808 BP.
XX
AC   AAD47110;
XX
DT   29-AUG-2003  (revised)
DT   27-JAN-2003  (first entry)
XX
```

```
DE   Mycobacterium sp. MTB72F-Leishmania sp. MAPS (aka r95f)-fusion DNA.
XX
KW   Vaccine; immunity; diagnostic agent; gene therapy; MTB72F; MAPS;
KW   chimeric; gene; ds.
XX
OS   Mycobacterium sp.
OS   Leishmania sp.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2796
FT                   /*tag=  a
FT                   /product= "Mycobacterium sp. MTB72F-Leishmania sp. MAPS
FT                   (aka r95f) fusion protein"
XX
PN   WO200272792-A2.
XX
PD   19-SEP-2002.
XX
PF   13-MAR-2002; 2002WO-US008223.
XX
PR   13-MAR-2001; 2001US-0275837P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Brannon M,  Guderian J;
XX
DR   WPI; 2002-759844/82.
DR   P-PSDB; AAE29731.
XX
PT   New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT   M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT   against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT   tuberculosis.
XX
PS   Example 6; Page 128-129; 155pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC   polynucleotide sequence encoding an antigen or an antigenic fragment from
CC   Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC   polypeptide or its fragment. The Leishmania polynucleotide is selected
CC   from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC   are used in methods for eliciting immune response in mammals. They are
CC   useful as vaccines to elicit protective immunity against pathogenic
CC   microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC   polypeptides are used for enhancing the expression of polynucleotides, as
CC   in vivo diagnostic agents and for raising antibodies in a non-human
CC   animal. The invention is used in gene therapy. The present sequence is
CC   Mycobacterium sp. MTB72F-Leishmania sp. thiol specific antioxidant [TSA;
CC   MAPS (aka r95f)] fusion DNA. This sequence comprises Mycobacterium sp.
CC   MTB72F (a 72 kDa poly-protein fusion construct comprising Ra12-TbH9-Ra35)
CC   linked to the Leishmania TSA (MAPS). (Updated on 29-AUG-2003 to
CC   standardise OS field)
XX
SQ   Sequence 2808 BP; 466 A; 889 C; 982 G; 471 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              7.36e-54      Length:          2808
Score:                  675.00        Matches:         132
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     6             Gaps:            0

US-09-684-215B-4 (1-132) x AAD47110 (1-2808)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 27
ADA26357
ID   ADA26357 standard; DNA; 2808 BP.
XX
AC   ADA26357;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB72F-MAPS (fusion r95F) protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2796
FT                   /*tag=  a
FT                   /product= "MTB72F-MAPS (fusion r95F)"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26364.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Disclosure; Fig 5; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
```

```
XX
SQ   Sequence 2808 BP; 466 A; 889 C; 982 G; 471 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                7.36e-54        Length:          2808
Score:                    675.00          Matches:         132
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       8               Gaps:            0

US-09-684-215B-4 (1-132) x ADA26357 (1-2808)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 28
ADA26355
ID   ADA26355 standard; DNA; 3030 BP.
XX
AC   ADA26355;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB-102F fusion protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .3030
FT                   /*tag=  a
FT                   /product= "MTB-120F fusion protein"
FT                   /note= "No stop codon given"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26356.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 5; Fig 3; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 3030 BP; 498 A; 966 C; 1062 G; 504 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                8.03e-54        Length:          3030
Score:                    675.00          Matches:         132
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       8               Gaps:            0

US-09-684-215B-4 (1-132) x ADA26355 (1-3030)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 654

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 714

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 774

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 834

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 894

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 954

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        955 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 990

RESULT 29
ADA26363
ID   ADA26363 standard; DNA; 3060 BP.
XX
AC   ADA26363;
XX
DT   20-NOV-2003  (first entry)
XX
DE   M. bovis MTB72F and 85b complex antigen (fusion MTB103F) DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine; MTB72F; 85B.
```

```
XX
OS   Chimeric.
OS   Mycobacterium bovis.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .3054
FT                   /*tag= a
FT                   /product= "MTB72F and 85b complex (fusion MTB103F)"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26370.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 84; Fig 11; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 3060 BP; 503 A; 977 C; 1070 G; 510 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                8.12e-54        Length:         3060
Score:                    675.00          Matches:        132
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       8               Gaps:           0

US-09-684-215B-4 (1-132) x ADA26363 (1-3060)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 30
ADA26362
ID   ADA26362 standard; DNA; 3104 BP.
XX
AC   ADA26362;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB72F-hTCC#1 (fusion MTB102tm2F) protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .3072
FT                   /*tag= a
FT                   /product= "MTB72F-hTCC#1 (fusion MTB102tm2F) protein"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26369.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 84; Fig 10; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 3104 BP; 510 A; 968 C; 1104 G; 522 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                8.25e-54        Length:         3104
Score:                    675.00          Matches:        132
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       8               Gaps:           0

US-09-684-215B-4 (1-132) x ADA26362 (1-3104)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

Search completed: April 29, 2004, 23:03:02
Job time : 382.474 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 22:43:07 ; Search time 483.789 Seconds
(without alignments)
1234.209 Million cell updates/sec

Title: US-09-684-215B-4
Perfect score: 675
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....QTKSGGTRTGNVTLAEGPPA 132

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061306_13266/app_query.fasta_1.1180
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=30 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09684215_@CGN_1_1_370_@runat_29042004_061306_13266
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	675	100.0	702	9	US-09-287-849-27	Sequence 27, Appl
2	675	100.0	702	15	US-10-359-460-27	Sequence 27, Appl
3	675	100.0	1002	15	US-10-098-732A-3	Sequence 3, Appli
4	675	100.0	1002	15	US-10-098-732A-5	Sequence 5, Appli
5	675	100.0	1068	9	US-09-712-363-15	Sequence 15, Appl
6	675	100.0	2181	16	US-10-369-983-1	Sequence 1, Appli
7	675	100.0	2190	15	US-10-098-732A-17	Sequence 17, Appl
8	675	100.0	2287	9	US-09-287-849-1	Sequence 1, Appli
9	675	100.0	2287	15	US-10-359-460-1	Sequence 1, Appli
10	675	100.0	2287	15	US-10-098-732A-15	Sequence 15, Appl
11	675	100.0	2287	16	US-10-359-459-1	Sequence 1, Appli
12	675	100.0	2451	16	US-10-369-983-8	Sequence 8, Appli
13	675	100.0	2487	16	US-10-369-983-7	Sequence 7, Appli
14	675	100.0	2637	16	US-10-369-983-6	Sequence 6, Appli
15	675	100.0	2808	15	US-10-098-732A-64	Sequence 64, Appl
16	675	100.0	2808	16	US-10-369-983-5	Sequence 5, Appli
17	675	100.0	3030	16	US-10-369-983-3	Sequence 3, Appli
18	675	100.0	3060	16	US-10-369-983-11	Sequence 11, Appl
19	675	100.0	3104	16	US-10-369-983-10	Sequence 10, Appl
20	675	100.0	3474	16	US-10-369-983-9	Sequence 9, Appli
21	670	99.3	447	15	US-10-193-002-4	Sequence 4, Appli
22	670	99.3	447	15	US-10-084-843-4	Sequence 4, Appli
23	670	99.3	447	15	US-10-098-732A-9	Sequence 9, Appli
24	670	99.3	1872	15	US-10-193-002-17	Sequence 17, Appl
25	670	99.3	1872	15	US-10-084-843-17	Sequence 17, Appl
26	670	99.3	1872	15	US-10-098-732A-1	Sequence 1, Appli
27	632	93.6	675	9	US-09-759-143-822	Sequence 822, App
28	632	93.6	675	9	US-09-780-669-822	Sequence 822, App
29	632	93.6	675	9	US-09-822-827-822	Sequence 822, App
30	632	93.6	675	9	US-09-895-793-822	Sequence 822, App
31	632	93.6	675	9	US-09-895-814-822	Sequence 822, App
32	632	93.6	675	14	US-10-012-896-822	Sequence 822, App
33	632	93.6	675	15	US-10-144-678A-822	Sequence 822, App
34	632	93.6	675	15	US-10-294-025-822	Sequence 822, App
35	632	93.6	822	9	US-09-736-457-1862	Sequence 1862, Ap
36	632	93.6	822	9	US-09-902-941-1862	Sequence 1862, Ap
37	632	93.6	822	9	US-09-849-626-1862	Sequence 1862, Ap
38	632	93.6	822	13	US-10-283-017-1862	Sequence 1862, Ap
39	632	93.6	822	15	US-10-017-754-1862	Sequence 1862, Ap
40	632	93.6	822	15	US-10-113-872-1862	Sequence 1862, Ap
41	632	93.6	894	13	US-09-878-722-236	Sequence 236, App
42	632	93.6	894	13	US-09-904-456-236	Sequence 236, App
43	632	93.6	900	9	US-09-735-705-353	Sequence 353, App
44	632	93.6	900	9	US-09-850-716A-353	Sequence 353, App
45	632	93.6	900	9	US-09-897-778-353	Sequence 353, App

ALIGNMENTS

RESULT 1
US-09-287-849-27
; Sequence 27, Application US/09287849
; Patent No. US20020009459A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556

; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein Ral2-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1
; NAME/KEY: CDS
; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ral2-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-09-287-849-27

Alignment Scores:

Pred. No.:	8.18e-71	Length:	702
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-4 (1-132) x US-09-287-849-27 (1-702)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 2
US-10-359-460-27
; Sequence 27, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein Ral2-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ral2-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-10-359-460-27

Alignment Scores:

Pred. No.:	8.18e-71	Length:	702
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-4 (1-132) x US-10-359-460-27 (1-702)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324
```

```
Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              |||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 3
US-10-098-732A-3
; Sequence 3, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 1002
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Ra35 mature
US-10-098-732A-3

Alignment Scores:
Pred. No.:                1.26e-70      Length:          1002
Score:                    675.00        Matches:         132
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       15            Gaps:            0

US-09-684-215B-4 (1-132) x US-10-098-732A-3 (1-1002)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              |||||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993

RESULT 4
US-10-098-732A-5
; Sequence 5, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1002
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Ra35FLMutSA
US-10-098-732A-5

Alignment Scores:
Pred. No.:                1.26e-70      Length:          1002
Score:                    675.00        Matches:         132
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       15            Gaps:            0

US-09-684-215B-4 (1-132) x US-10-098-732A-5 (1-1002)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              |||||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993

RESULT 5
US-09-712-363-15
; Sequence 15, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
;  APPLICANT: Eisenberg, David
;  APPLICANT:  Rotstein, Sergio H.
;  APPLICANT:  Marcotte, Edward M.
;  TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
;  TITLE OF INVENTION:  INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
```

```
;  FILE REFERENCE: 07419-032001
;  CURRENT APPLICATION NUMBER: US/09/712,363
;  CURRENT FILING DATE:  2000-11-13
;  PRIOR APPLICATION NUMBER: PCT/US00/02246
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: 60/179,531
;  PRIOR FILING DATE: 2000-02-01
;  PRIOR APPLICATION NUMBER: 60/117,844
;  PRIOR FILING DATE: 1999-01-29
;  PRIOR APPLICATION NUMBER: 60/118,206,
;  PRIOR FILING DATE: 1999-02-01
;  PRIOR APPLICATION NUMBER: 60/126,593
;  PRIOR FILING DATE: 1999-03-26
;  PRIOR APPLICATION NUMBER: 60/134,093
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/134,092
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/165,124
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/165,086
;  PRIOR FILING DATE: 1999-11-12
;  NUMBER OF SEQ ID NOS: 292
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 1068
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
US-09-712-363-15

Alignment Scores:
Pred. No.:                1.36e-70        Length:          1068
Score:                    675.00          Matches:         132
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-09-684-215B-4 (1-132) x US-09-712-363-15 (1-1068)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 729

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 789

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 849

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 909

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 969

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1029

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db       1030 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1065

RESULT 6
US-10-369-983-1
; Sequence 1, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT: Guderian, Jeff
;  APPLICANT: Reed, Steven
;  APPLICANT: Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2181
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:mutated
;   OTHER INFORMATION: MTB32-MTB39F fusion protein (MTB32MutSA)
US-10-369-983-1

Alignment Scores:
Pred. No.:                3.22e-70        Length:          2181
Score:                    675.00          Matches:         132
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       16              Gaps:            0

US-09-684-215B-4 (1-132) x US-10-369-983-1 (1-2181)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 993

RESULT 7
US-10-098-732A-17
; Sequence 17, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
```

```
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 2190
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:MTB72FMutSA
;   OTHER INFORMATION: (Ra12-TbH9-Ra35MutSA)
US-10-098-732A-17

Alignment Scores:
Pred. No.:                3.23e-70        Length:          2190
Score:                    675.00          Matches:         132
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       15              Gaps:            0

US-09-684-215B-4 (1-132) x US-10-098-732A-17 (1-2190)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 141

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 201

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 321

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 381

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        382 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 417

RESULT 8
US-09-287-849-1
; Sequence 1, Application US/09287849
; Patent No. US20020009459A1
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Alderson, Mark
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION:  Fusion Protiens of Mycobacterium tuberculosis Antigens
;  TITLE OF INVENTION:  and Their Uses
;  FILE REFERENCE:  014058-009020US
;  CURRENT APPLICATION NUMBER: US/09/287,849
;  CURRENT FILING DATE:  1999-04-07
;  PRIOR APPLICATION NUMBER:  US 08/818,112
;  PRIOR FILING DATE:  1997-03-13
;  PRIOR APPLICATION NUMBER:  US 08/942,578
;  PRIOR FILING DATE:  1997-10-01
;  PRIOR APPLICATION NUMBER:  US 09/025,197
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 09/056,556
;  PRIOR FILING DATE:  1998-04-07
;  PRIOR APPLICATION NUMBER:  US 09/223,040
;  PRIOR FILING DATE:  1998-12-30
;  NUMBER OF SEQ ID NOS:  46
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2287
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
;   OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39
;   OTHER INFORMATION: fusion)
;   NAME/KEY: modified_base
;   LOCATION: (30)
;   OTHER INFORMATION: n = g, a, c or t
;   NAME/KEY: modified_base
;   LOCATION: (33)
;   OTHER INFORMATION: n = g, a, c or t
;   NAME/KEY: CDS
;   LOCATION: (42)..(2231)
;   NAME/KEY: modified_base
;   LOCATION: (2270)
;   OTHER INFORMATION: n = g, a, c or t
US-09-287-849-1

Alignment Scores:
Pred. No.:                3.41e-70        Length:          2287
Score:                    675.00          Matches:         132
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-09-684-215B-4 (1-132) x US-09-287-849-1 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458

RESULT 9
US-10-359-460-1
; Sequence 1, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Alderson, Mark
```

```
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
;  TITLE OF INVENTION:  and Their Uses
;  FILE REFERENCE: 014058-009020US
;  CURRENT APPLICATION NUMBER: US/10/359,460
;  CURRENT FILING DATE:  2003-02-05
;  PRIOR APPLICATION NUMBER: US/09/287,849
;  PRIOR FILING DATE: 1999-04-07
;  PRIOR APPLICATION NUMBER: US 08/818,112
;  PRIOR FILING DATE: 1997-03-13
;  PRIOR APPLICATION NUMBER: US 08/942,578
;  PRIOR FILING DATE: 1997-10-01
;  PRIOR APPLICATION NUMBER: US 09/025,197
;  PRIOR FILING DATE: 1998-02-18
;  PRIOR APPLICATION NUMBER: US 09/056,556
;  PRIOR FILING DATE: 1998-04-07
;  PRIOR APPLICATION NUMBER: US 09/223,040
;  PRIOR FILING DATE: 1998-12-30
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2287
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
;   OTHER INFORMATION: protein Ral2-TbH9-Ra35 (designated Mtb32-Mtb39
;   OTHER INFORMATION: fusion)
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (30)
;   OTHER INFORMATION: n = g, a, c or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (33)
;   OTHER INFORMATION: n = g, a, c or t
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (42)..(2231)
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (2270)
;   OTHER INFORMATION: n = g, a, c or t
US-10-359-460-1

Alignment Scores:
Pred. No.:               3.41e-70      Length:          2287
Score:                   675.00        Matches:         132
Percent Similarity:      100.00%       Conservative:    0
Best Local Similarity:   100.00%       Mismatches:      0
Query Match:             100.00%       Indels:          0
DB:                      15            Gaps:            0

US-09-684-215B-4 (1-132) x US-10-359-460-1 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458

RESULT 10
US-10-098-732A-15
; Sequence 15, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 2287
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
;   OTHER INFORMATION: protein MTB72F (Ral2-TbH9-Ra35 or MTB32-MTB39
;   OTHER INFORMATION: fusion)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (42)..(2231)
;   OTHER INFORMATION: MTB72F
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1)..(2287)
;   OTHER INFORMATION: n = g, a, c or t
US-10-098-732A-15

Alignment Scores:
Pred. No.:               3.41e-70      Length:          2287
Score:                   675.00        Matches:         132
Percent Similarity:      100.00%       Conservative:    0
Best Local Similarity:   100.00%       Mismatches:      0
Query Match:             100.00%       Indels:          0
DB:                      15            Gaps:            0

US-09-684-215B-4 (1-132) x US-10-098-732A-15 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302
```

```
Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy         121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
               ||||||||||||||||||||||||||||||||||||
Db         423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458

RESULT 11
US-10-359-459-1
; Sequence 1, Application US/10359459
; Publication No. US20040013677A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Alderson, Mark
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
;  TITLE OF INVENTION:  and Their Uses
;  FILE REFERENCE: 014058-009010US
;  CURRENT APPLICATION NUMBER: US/10/359,459
;  CURRENT FILING DATE:  2003-02-05
;  PRIOR APPLICATION NUMBER: US/09/223,040
;  PRIOR FILING DATE: 1998-12-30
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2287
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
;   OTHER INFORMATION: protein Ral2-TbH9-Ra35
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (30)
;   OTHER INFORMATION: n = g, a, c or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (33)
;   OTHER INFORMATION: n = g, a, c or t
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (42)..(2231)
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (2270)
;   OTHER INFORMATION: n = g, a, c or t
US-10-359-459-1

Alignment Scores:
Pred. No.:              3.41e-70        Length:         2287
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     16              Gaps:           0

US-09-684-215B-4 (1-132) x US-10-359-459-1 (1-2287)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy         121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
               ||||||||||||||||||||||||||||||||||||
Db         423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458

RESULT 12
US-10-369-983-8
; Sequence 8, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 2451
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB81F (MTB72F-DPV)
US-10-369-983-8

Alignment Scores:
Pred. No.:              3.71e-70        Length:         2451
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     16              Gaps:           0

US-09-684-215B-4 (1-132) x US-10-369-983-8 (1-2451)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 13
US-10-369-983-7
; Sequence 7, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 2487
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB83F (MTB72F-MTI)
US-10-369-983-7

Alignment Scores:
Pred. No.:                3.77e-70        Length:         2487
Score:                    675.00          Matches:        132
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       16              Gaps:           0

US-09-684-215B-4 (1-132) x US-10-369-983-7 (1-2487)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 14
US-10-369-983-6
; Sequence 6, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 2637
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB89F (MTB72F-Erd14)
US-10-369-983-6

Alignment Scores:
Pred. No.:                4.05e-70        Length:         2637
Score:                    675.00          Matches:        132
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       16              Gaps:           0

US-09-684-215B-4 (1-132) x US-10-369-983-6 (1-2637)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 15
US-10-098-732A-64
; Sequence 64, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
```

```
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
;   LENGTH: 2808
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:MTB72F-MAPS
;   OTHER INFORMATION: (r95f) fusion construct, TB MTB72F (Ra12-TbH9-Ra35)
;   OTHER INFORMATION: linked to Leishmania thiol-specific antioxidant
;   OTHER INFORMATION: (TSA or MAPS)
US-10-098-732A-64

Alignment Scores:
Pred. No.:              4.37e-70        Length:         2808
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     15              Gaps:           0

US-09-684-215B-4 (1-132) x US-10-098-732A-64 (1-2808)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 16
US-10-369-983-5
; Sequence 5, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 2808
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: R95F (MTB72F-MAPS)
US-10-369-983-5

Alignment Scores:
Pred. No.:              4.37e-70        Length:         2808
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     16              Gaps:           0

US-09-684-215B-4 (1-132) x US-10-369-983-5 (1-2808)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 17
US-10-369-983-3
; Sequence 3, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 3030
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:MTB-102F fusion
```

```
;   OTHER INFORMATION: protein
US-10-369-983-3

Alignment Scores:
Pred. No.:              4.79e-70        Length:         3030
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     16              Gaps:           0

US-09-684-215B-4 (1-132) x US-10-369-983-3 (1-3030)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 654

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 714

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 774

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 834

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 894

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 954

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        955 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 990

RESULT 18
US-10-369-983-11
; Sequence 11, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 3060
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB103F (MTB72F-85b)
US-10-369-983-11

Alignment Scores:
Pred. No.:              4.84e-70        Length:         3060
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     16              Gaps:           0

US-09-684-215B-4 (1-132) x US-10-369-983-11 (1-3060)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 19
US-10-369-983-10
; Sequence 10, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 3104
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB102tm2F (MTB102FTM, MTB72F-hTCC#1)
US-10-369-983-10

Alignment Scores:
Pred. No.:              4.93e-70        Length:         3104
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     16              Gaps:           0

US-09-684-215B-4 (1-132) x US-10-369-983-10 (1-3104)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy         121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
                    ||||||||||||||||||||||||||||||||
Db         385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 20
US-10-369-983-9
; Sequence 9, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 3474
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB114F (MTB72F-mTCC#2)
US-10-369-983-9

Alignment Scores:
Pred. No.:              5.65e-70        Length:         3474
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     16              Gaps:           0

US-09-684-215B-4 (1-132) x US-10-369-983-9 (1-3474)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy         121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
                ||||||||||||||||||||||||||||||||
Db         385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 21
US-10-193-002-4
; Sequence 4, Application US/10193002
; Publication No. US20030135026A1
;  GENERAL INFORMATION:
;       APPLICANT: Reed, Steven G.
;                  Skeiky, Yasir A.W.
;                  Dillon, Davin C.
;                  Campos-Neto, Antonia
;                  Houghton, Raymond
;                  Vedvick, Thomas S.
;                  Twardzik, Daniel R.
;                  Lodes, Michael J.
;                  Hendrickson, Ronald C.
;       TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                           TUBERCULOSIS
;       NUMBER OF SEQUENCES: 350
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SEED and BERRY LLP
;            STREET: 6300 Columbia Center, 701 Fifth Avenue
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: USA
;            ZIP: 98104-7092
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/193,002
;            FILING DATE: 10-Jul-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/072,596
;            FILING DATE: 05-MAY-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Maki, David J.
;            REGISTRATION NUMBER: 31,392
;            REFERENCE/DOCKET NUMBER: 210121.417C9
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (206) 622-4900
;            TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 447 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-193-002-4

Alignment Scores:
Pred. No.:              1.86e-70        Length:         447
Score:                  670.00          Matches:        131
Percent Similarity:     99.24%          Conservative:   0
Best Local Similarity:  99.24%          Mismatches:     1
Query Match:            99.26%          Indels:         0
DB:                     15              Gaps:           0
```

```
US-09-684-215B-4 (1-132) x US-10-193-002-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 22
US-10-084-843-4
; Sequence 4, Application US/10084843
; Publication No. US20030143243A1
;   GENERAL INFORMATION:
;        APPLICANT: Reed, Steven G.
;                   Skeiky, Yasir A.W.
;                   Dillon, Davin C.
;                   Campos-Neto, Antonio
;                   Houghton, Raymond
;                   Vedvick, Thomas S.
;                   Twardzik, Daniel R.
;                   Lodes, Michael J.
;                   Hendrickson, Ronald C.
;        TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;                            AND DIAGNOSIS OF TUBERCULOSIS
;        NUMBER OF SEQUENCES: 355
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SEED and BERRY LLP
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/084,843
;             FILING DATE: 25-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/072,967
;             FILING DATE: 05-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Maki, David J.
;             REGISTRATION NUMBER: 31,392
;             REFERENCE/DOCKET NUMBER: 210121.411C9
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 447 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-084-843-4

Alignment Scores:
Pred. No.:                1.86e-70        Length:          447
Score:                    670.00          Matches:         131
Percent Similarity:       99.24%          Conservative:    0
Best Local Similarity:    99.24%          Mismatches:      1
Query Match:              99.26%          Indels:          0
DB:                       15              Gaps:            0

US-09-684-215B-4 (1-132) x US-10-084-843-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 23
US-10-098-732A-9
; Sequence 9, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 447
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:MTBRa12 (MTB32A
;   OTHER INFORMATION: C-terminus)
```

US-10-098-732A-9

Alignment Scores:

Pred. No.:	1.86e-70	Length:	447
Score:	670.00	Matches:	131
Percent Similarity:	99.24%	Conservative:	0
Best Local Similarity:	99.24%	Mismatches:	1
Query Match:	99.26%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-4 (1-132) x US-10-098-732A-9 (1-447)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

```
RESULT 24
US-10-193-002-17
; Sequence 17, Application US/10193002
; Publication No. US20030135026A1
;   GENERAL INFORMATION:
;        APPLICANT: Reed, Steven G.
;                   Skeiky, Yasir A.W.
;                   Dillon, Davin C.
;                   Campos-Neto, Antonia
;                   Houghton, Raymond
;                   Vedvick, Thomas S.
;                   Twardzik, Daniel R.
;                   Lodes, Michael J.
;                   Hendrickson, Ronald C.
;        TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                            TUBERCULOSIS
;        NUMBER OF SEQUENCES: 350
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SEED and BERRY LLP
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/193,002
;             FILING DATE: 10-Jul-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/072,596
;             FILING DATE: 05-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Maki, David J.
;             REGISTRATION NUMBER: 31,392
;             REFERENCE/DOCKET NUMBER: 210121.417C9
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 17:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1872 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-193-002-17

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 1.05e-69 | Length: | 1872 |
| Score: | 670.00 | Matches: | 131 |
| Percent Similarity: | 99.24% | Conservative: | 0 |
| Best Local Similarity: | 99.24% | Mismatches: | 1 |
| Query Match: | 99.26% | Indels: | 0 |
| DB: | 15 | Gaps: | 0 |

US-09-684-215B-4 (1-132) x US-10-193-002-17 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153

```
RESULT 25
US-10-084-843-17
; Sequence 17, Application US/10084843
; Publication No. US20030143243A1
;   GENERAL INFORMATION:
;   APPLICANT: Reed, Steven G.
;   Skeiky, Yasir A.W.
;   Dillon, Davin C.
;   Campos-Neto, Antonio
;   Houghton, Raymond
;   Vedvick, Thomas S.
;   Twardzik, Daniel R.
;   Lodes, Michael J.
;   Hendrickson, Ronald C.
;   TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;   AND DIAGNOSIS OF TUBERCULOSIS
```
```

```
;       NUMBER OF SEQUENCES: 355
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SEED and BERRY LLP
;            STREET: 6300 Columbia Center, 701 Fifth Avenue
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: USA
;            ZIP: 98104-7092
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/084,843
;            FILING DATE: 25-Feb-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/072,967
;            FILING DATE: 05-MAY-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Maki, David J.
;            REGISTRATION NUMBER: 31,392
;            REFERENCE/DOCKET NUMBER: 210121.411C9
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (206) 622-4900
;            TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 17:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1872 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-084-843-17

Alignment Scores:
Pred. No.:              1.05e-69        Length:         1872
Score:                  670.00          Matches:        131
Percent Similarity:     99.24%          Conservative:   0
Best Local Similarity:  99.24%          Mismatches:     1
Query Match:            99.26%          Indels:         0
DB:                     15              Gaps:           0

US-09-684-215B-4 (1-132) x US-10-084-843-17 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153
```

```
RESULT 26
US-10-098-732A-1
; Sequence 1, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1872
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: MTB32A (Ra35FL)
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1)..(1872)
;   OTHER INFORMATION: n = g, a, c or t
US-10-098-732A-1

Alignment Scores:
Pred. No.:              1.05e-69        Length:         1872
Score:                  670.00          Matches:        131
Percent Similarity:     99.24%          Conservative:   0
Best Local Similarity:  99.24%          Mismatches:     1
Query Match:            99.26%          Indels:         0
DB:                     15              Gaps:           0

US-09-684-215B-4 (1-132) x US-10-098-732A-1 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153

RESULT 27
US-09-759-143-822
; Sequence 822, Application US/09759143
; Patent No. US20020022248A1
```

```
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-143-822

Alignment Scores:
Pred. No.:              9.88e-66        Length:         675
Score:                  632.00          Matches:        126
Percent Similarity:     96.21%          Conservative:   1
Best Local Similarity:  95.45%          Mismatches:     1
Query Match:            93.63%          Indels:         4
DB:                     9               Gaps:           1

US-09-684-215B-4 (1-132) x US-09-759-143-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::              ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405

RESULT 28
US-09-780-669-822
; Sequence 822, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C24
;  CURRENT APPLICATION NUMBER: US/09/780,669
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS:  943
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-780-669-822

Alignment Scores:
Pred. No.:              9.88e-66        Length:         675
Score:                  632.00          Matches:        126
Percent Similarity:     96.21%          Conservative:   1
Best Local Similarity:  95.45%          Mismatches:     1
Query Match:            93.63%          Indels:         4
DB:                     9               Gaps:           1

US-09-684-215B-4 (1-132) x US-09-780-669-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::              ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405

RESULT 29
```

```
US-09-822-827-822
; Sequence 822, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C1
;  CURRENT APPLICATION NUMBER: US/09/822,827
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-822-827-822

Alignment Scores:
Pred. No.:              9.88e-66      Length:         675
Score:                  632.00        Matches:        126
Percent Similarity:     96.21%        Conservative:   1
Best Local Similarity:  95.45%        Mismatches:     1
Query Match:            93.63%        Indels:         4
DB:                     9             Gaps:           1

US-09-684-215B-4 (1-132) x US-09-822-827-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||:::           ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405

RESULT 30
US-09-895-793-822
; Sequence 822, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-793-822

Alignment Scores:
Pred. No.:              9.88e-66      Length:         675
Score:                  632.00        Matches:        126
Percent Similarity:     96.21%        Conservative:   1
Best Local Similarity:  95.45%        Mismatches:     1
Query Match:            93.63%        Indels:         4
DB:                     9             Gaps:           1

US-09-684-215B-4 (1-132) x US-09-895-793-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||:::           ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405


Search completed: April 30, 2004, 05:10:21
Job time : 493.039 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 22:28:57 ; Search time 67.5789 Seconds
(without alignments)
1083.969 Million cell updates/sec

Title: US-09-684-215B-4
Perfect score: 675
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....QTKSGGTRTGNVTLAEGPPA 132

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061306_13249/app_query.fasta_1.1180
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_115_@runat_29042004_061306_13249 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	675	100.0	702	4	US-09-287-849-27	Sequence 27, Appl
2	675	100.0	2287	4	US-09-223-040-1	Sequence 1, Appli
3	675	100.0	2287	4	US-09-287-849-1	Sequence 1, Appli
4	675	100.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
5	675	100.0	4411529	3	US-09-103-840A-1	Sequence 1, Appli
6	670	99.3	447	3	US-08-818-112-4	Sequence 4, Appli
7	670	99.3	447	4	US-08-818-111-4	Sequence 4, Appli
8	670	99.3	447	4	US-09-056-556-4	Sequence 4, Appli
9	670	99.3	447	4	US-09-072-596-4	Sequence 4, Appli
10	670	99.3	447	4	US-09-072-967-4	Sequence 4, Appli
11	670	99.3	1872	3	US-08-818-112-17	Sequence 17, Appl
12	670	99.3	1872	4	US-08-818-111-17	Sequence 17, Appl
13	670	99.3	1872	4	US-09-056-556-17	Sequence 17, Appl
14	670	99.3	1872	4	US-09-072-596-17	Sequence 17, Appl
15	670	99.3	1872	4	US-09-072-967-17	Sequence 17, Appl
16	632	93.6	675	4	US-09-636-215-822	Sequence 822, App
17	632	93.6	675	4	US-09-685-166A-822	Sequence 822, App
18	632	93.6	822	4	US-09-736-457-1862	Sequence 1862, Ap
19	632	93.6	900	4	US-09-643-597-353	Sequence 353, App
20	632	93.6	900	4	US-09-606-421B-353	Sequence 353, App
21	632	93.6	915	4	US-09-636-215-834	Sequence 834, App
22	632	93.6	915	4	US-09-685-166A-834	Sequence 834, App
23	632	93.6	945	4	US-09-736-457-1861	Sequence 1861, Ap
24	632	93.6	1012	4	US-09-643-597-351	Sequence 351, App
25	632	93.6	1012	4	US-09-606-421B-351	Sequence 351, App
26	632	93.6	1203	4	US-09-636-215-851	Sequence 851, App
27	632	93.6	1203	4	US-09-685-166A-851	Sequence 851, App
28	632	93.6	1464	4	US-09-620-412C-348	Sequence 348, App
29	632	93.6	1464	4	US-09-598-419-348	Sequence 348, App
30	632	93.6	1557	4	US-09-620-412C-332	Sequence 332, App
31	632	93.6	1557	4	US-09-598-419-332	Sequence 332, App
32	632	93.6	1578	4	US-09-556-877-188	Sequence 188, App
33	632	93.6	1578	4	US-09-620-412C-188	Sequence 188, App
34	632	93.6	1578	4	US-09-598-419-188	Sequence 188, App
35	632	93.6	1752	4	US-09-620-412C-352	Sequence 352, App
36	632	93.6	1752	4	US-09-598-419-352	Sequence 352, App
37	632	93.6	1758	4	US-09-620-412C-336	Sequence 336, App
38	632	93.6	1758	4	US-09-598-419-336	Sequence 336, App
39	632	93.6	1860	4	US-09-620-412C-308	Sequence 308, App
40	632	93.6	1860	4	US-09-598-419-308	Sequence 308, App
41	632	93.6	1896	4	US-09-620-412C-324	Sequence 324, App
42	632	93.6	1896	4	US-09-598-419-324	Sequence 324, App
43	632	93.6	1941	4	US-09-620-412C-316	Sequence 316, App
44	632	93.6	1941	4	US-09-598-419-316	Sequence 316, App
45	632	93.6	1965	4	US-09-620-412C-340	Sequence 340, App

ALIGNMENTS

RESULT 1
US-09-287-849-27
; Sequence 27, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1

; NAME/KEY: CDS
; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-09-287-849-27

Alignment Scores:

Pred. No.:	2.07e-65	Length:	702
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-4 (1-132) x US-09-287-849-27 (1-702)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420
```

RESULT 2
US-09-223-040-1
; Sequence 1, Application US/09223040
; Patent No. 6544522
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009010US
; CURRENT APPLICATION NUMBER: US/09/223,040
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
; OTHER INFORMATION: protein Ra12-TbH9-Ra35
; NAME/KEY: modified_base
; LOCATION: (30)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: modified_base
; LOCATION: (33)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
; NAME/KEY: modified_base
; LOCATION: (2270)
; OTHER INFORMATION: n = g, a, c or t
US-09-223-040-1

Alignment Scores:

Pred. No.:	9.78e-65	Length:	2287
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-4 (1-132) x US-09-223-040-1 (1-2287)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

RESULT 3
US-09-287-849-1
; Sequence 1, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
; OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39
; OTHER INFORMATION: fusion)
; NAME/KEY: modified_base
; LOCATION: (30)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: modified_base
; LOCATION: (33)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
; NAME/KEY: modified_base
; LOCATION: (2270)
; OTHER INFORMATION: n = g, a, c or t
US-09-287-849-1

Alignment Scores:

Pred. No.:	9.78e-65	Length:	2287
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-4 (1-132) x US-09-287-849-1 (1-2287)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

RESULT 4
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:

Pred. No.:	2.03e-60	Length:	4403765
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-09-684-215B-4 (1-132) x US-09-103-840A-2 (1-4403765)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151984 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 152043

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152044 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 152103

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152104 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 152163

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152164 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 152223

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152224 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 152283

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152284 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 152343

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||
Db     152344 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 152379
```

RESULT 5
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1

```
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.:                2.04e-60        Length:         4411529
Score:                    675.00          Matches:        132
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       3               Gaps:           0

US-09-684-215B-4 (1-132) x US-09-103-840A-1 (1-4411529)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151815 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 151874

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151875 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 151934

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151935 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 151994

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151995 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 152054

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152055 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 152114

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152115 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 152174

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db     152175 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 152210

RESULT 6
US-08-818-112-4
; Sequence 4, Application US/08818112
; Patent No. 6290969
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  153
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,112
;      FILING DATE:  13-MAR-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-818-112-4

Alignment Scores:
Pred. No.:                4.05e-65        Length:         447
Score:                    670.00          Matches:        131
Percent Similarity:       99.24%          Conservative:   0
Best Local Similarity:    99.24%          Mismatches:     1
Query Match:              99.26%          Indels:         0
DB:                       3               Gaps:           0

US-09-684-215B-4 (1-132) x US-08-818-112-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 7
US-08-818-111-4
; Sequence 4, Application US/08818111
; Patent No. 6338852
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  148
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
```

```
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,111
;      FILING DATE:  13-MAR-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-818-111-4

Alignment Scores:
Pred. No.:                4.05e-65        Length:          447
Score:                    670.00          Matches:         131
Percent Similarity:       99.24%          Conservative:    0
Best Local Similarity:    99.24%          Mismatches:      1
Query Match:              99.26%          Indels:          0
DB:                       4               Gaps:            0

US-09-684-215B-4 (1-132) x US-08-818-111-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 8
US-09-056-556-4
; Sequence 4, Application US/09056556
; Patent No. 6350456
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND TREAT
;    NUMBER OF SEQUENCES:  241
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,556
;      FILING DATE:  07-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.457
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-056-556-4

Alignment Scores:
Pred. No.:                4.05e-65        Length:          447
Score:                    670.00          Matches:         131
Percent Similarity:       99.24%          Conservative:    0
Best Local Similarity:    99.24%          Mismatches:      1
Query Match:              99.26%          Indels:          0
DB:                       4               Gaps:            0

US-09-684-215B-4 (1-132) x US-09-056-556-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 9
US-09-072-596-4
```

```
; Sequence 4, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-4

Alignment Scores:
Pred. No.:                4.05e-65        Length:         447
Score:                    670.00          Matches:        131
Percent Similarity:       99.24%          Conservative:   0
Best Local Similarity:    99.24%          Mismatches:     1
Query Match:              99.26%          Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-4 (1-132) x US-09-072-596-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 10
US-09-072-967-4
; Sequence 4, Application US/09072967
; Patent No. 6592877
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  355
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,967
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-967-4

Alignment Scores:
Pred. No.:                4.05e-65        Length:         447
Score:                    670.00          Matches:        131
Percent Similarity:       99.24%          Conservative:   0
Best Local Similarity:    99.24%          Mismatches:     1
Query Match:              99.26%          Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-4 (1-132) x US-09-072-967-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
                ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db         311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy         121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
                ||||||||||||||||||||||||||||||||||||
Db         371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 11
US-08-818-112-17
; Sequence 17, Application US/08818112
; Patent No. 6290969
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  153
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,112
;      FILING DATE:  13-MAR-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-818-112-17

Alignment Scores:
Pred. No.:              2.66e-64        Length:         1872
Score:                  670.00          Matches:        131
Percent Similarity:     99.24%          Conservative:   0
Best Local Similarity:  99.24%          Mismatches:     1
Query Match:            99.26%          Indels:         0
DB:                     3               Gaps:           0

US-09-684-215B-4 (1-132) x US-08-818-112-17 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153

RESULT 12
US-08-818-111-17
; Sequence 17, Application US/08818111
; Patent No. 6338852
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  148
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,111
;      FILING DATE:  13-MAR-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  1872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-818-111-17

Alignment Scores:
Pred. No.:                2.66e-64        Length:          1872
Score:                    670.00          Matches:         131
Percent Similarity:       99.24%          Conservative:    0
Best Local Similarity:    99.24%          Mismatches:      1
Query Match:              99.26%          Indels:          0
DB:                       4               Gaps:            0

US-09-684-215B-4 (1-132) x US-08-818-111-17 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              |||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153

RESULT 13
US-09-056-556-17
; Sequence 17, Application US/09056556
; Patent No. 6350456
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND          TREATM
;    NUMBER OF SEQUENCES:  241
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,556
;      FILING DATE:  07-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.457
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-056-556-17

Alignment Scores:
Pred. No.:                2.66e-64        Length:          1872
Score:                    670.00          Matches:         131
Percent Similarity:       99.24%          Conservative:    0
Best Local Similarity:    99.24%          Mismatches:      1
Query Match:              99.26%          Indels:          0
DB:                       4               Gaps:            0

US-09-684-215B-4 (1-132) x US-09-056-556-17 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              |||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153

RESULT 14
US-09-072-596-17
; Sequence 17, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-17

Alignment Scores:
Pred. No.:                2.66e-64        Length:          1872
Score:                    670.00          Matches:         131
Percent Similarity:       99.24%          Conservative:    0
Best Local Similarity:    99.24%          Mismatches:      1
Query Match:              99.26%          Indels:          0
DB:                       4               Gaps:            0

US-09-684-215B-4 (1-132) x US-09-072-596-17 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153

RESULT 15
US-09-072-967-17
; Sequence 17, Application US/09072967
; Patent No. 6592877
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  355
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,967
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-967-17

Alignment Scores:
Pred. No.:                2.66e-64        Length:          1872
Score:                    670.00          Matches:         131
Percent Similarity:       99.24%          Conservative:    0
Best Local Similarity:    99.24%          Mismatches:      1
Query Match:              99.26%          Indels:          0
DB:                       4               Gaps:            0

US-09-684-215B-4 (1-132) x US-09-072-967-17 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
```

```
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153

RESULT 16
US-09-636-215-822
; Sequence 822, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.42717C17
;  CURRENT APPLICATION NUMBER: US/09/636,215
;  CURRENT FILING DATE:  2000-08-10
;  NUMBER OF SEQ ID NOS:  852
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-636-215-822

Alignment Scores:
Pred. No.:                1.03e-60        Length:         675
Score:                    632.00          Matches:        126
Percent Similarity:       96.21%          Conservative:   1
Best Local Similarity:    95.45%          Mismatches:     1
Query Match:              93.63%          Indels:         4
DB:                       4               Gaps:           1

US-09-684-215B-4 (1-132) x US-09-636-215-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::             ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
```

```
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405

RESULT 17
US-09-685-166A-822
; Sequence 822, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C21
;  CURRENT APPLICATION NUMBER: US/09/685,166A
;  CURRENT FILING DATE:  2000-10-10
;  NUMBER OF SEQ ID NOS: 898
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-685-166A-822

Alignment Scores:
Pred. No.:                1.03e-60        Length:         675
Score:                    632.00          Matches:        126
Percent Similarity:       96.21%          Conservative:   1
Best Local Similarity:    95.45%          Mismatches:     1
Query Match:              93.63%          Indels:         4
DB:                       4               Gaps:           1

US-09-684-215B-4 (1-132) x US-09-685-166A-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::             ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
```

```
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405

RESULT 18
US-09-736-457-1862
; Sequence 1862, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1862
;   LENGTH: 822
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-736-457-1862

Alignment Scores:
Pred. No.:                1.34e-60        Length:          822
Score:                    632.00          Matches:         126
Percent Similarity:       96.21%          Conservative:    1
Best Local Similarity:    95.45%          Mismatches:      1
Query Match:              93.63%          Indels:          4
DB:                       4               Gaps:            1

US-09-684-215B-4 (1-132) x US-09-736-457-1862 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405

RESULT 19
US-09-643-597-353
; Sequence 353, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
;  CURRENT APPLICATION NUMBER: US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 353
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-643-597-353

Alignment Scores:
Pred. No.:                1.51e-60        Length:          900
Score:                    632.00          Matches:         126
Percent Similarity:       96.21%          Conservative:    1
Best Local Similarity:    95.45%          Mismatches:      1
Query Match:              93.63%          Indels:          4
DB:                       4               Gaps:            1

US-09-684-215B-4 (1-132) x US-09-643-597-353 (1-900)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405

RESULT 20
US-09-606-421B-353
; Sequence 353, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
```

; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-421B-353

Alignment Scores:

Pred. No.:	1.51e-60	Length:	900
Score:	632.00	Matches:	126
Percent Similarity:	96.21%	Conservative:	1
Best Local Similarity:	95.45%	Mismatches:	1
Query Match:	93.63%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-4 (1-132) x US-09-606-421B-353 (1-900)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::           ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 21
US-09-636-215-834
; Sequence 834, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel

; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-636-215-834

Alignment Scores:

Pred. No.:	1.54e-60	Length:	915
Score:	632.00	Matches:	126
Percent Similarity:	96.21%	Conservative:	1
Best Local Similarity:	95.45%	Mismatches:	1
Query Match:	93.63%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-4 (1-132) x US-09-636-215-834 (1-915)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::           ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 22
US-09-685-166A-834
; Sequence 834, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel

; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-834

Alignment Scores:

Pred. No.:	1.54e-60	Length:	915
Score:	632.00	Matches:	126
Percent Similarity:	96.21%	Conservative:	1
Best Local Similarity:	95.45%	Mismatches:	1
Query Match:	93.63%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-4 (1-132) x US-09-685-166A-834 (1-915)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 23
US-09-736-457-1861
; Sequence 1861, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1861
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-1861

Alignment Scores:

Pred. No.:	1.61e-60	Length:	945
Score:	632.00	Matches:	126
Percent Similarity:	96.21%	Conservative:	1
Best Local Similarity:	95.45%	Mismatches:	1
Query Match:	93.63%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-4 (1-132) x US-09-736-457-1861 (1-945)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 24
US-09-643-597-351
; Sequence 351, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 351
; LENGTH: 1012
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-643-597-351

Alignment Scores:

Pred. No.:	1.76e-60	Length:	1012
Score:	632.00	Matches:	126
Percent Similarity:	96.21%	Conservative:	1
Best Local Similarity:	95.45%	Mismatches:	1
Query Match:	93.63%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-4 (1-132) x US-09-643-597-351 (1-1012)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::                ||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 174

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 234

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 294

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 354

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 414

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||
Db        415 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 450
```

RESULT 25
US-09-606-421B-351
; Sequence 351, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 351
; LENGTH: 1012
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-421B-351

Alignment Scores:

Pred. No.:	1.76e-60	Length:	1012
Score:	632.00	Matches:	126
Percent Similarity:	96.21%	Conservative:	1
Best Local Similarity:	95.45%	Mismatches:	1
Query Match:	93.63%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-4 (1-132) x US-09-606-421B-351 (1-1012)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::                ||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 174

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 234

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 294

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 354

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 414

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||
Db        415 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 450
```

RESULT 26
US-09-636-215-851
; Sequence 851, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-636-215-851

Alignment Scores:

Pred. No.:	2.21e-60	Length:	1203
Score:	632.00	Matches:	126
Percent Similarity:	96.21%	Conservative:	1
Best Local Similarity:	95.45%	Mismatches:	1

```
Query Match:          93.63%      Indels:        4
DB:                   4           Gaps:          1

US-09-684-215B-4 (1-132) x US-09-636-215-851 (1-1203)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::            ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405

RESULT 27
US-09-685-166A-851
; Sequence 851, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C21
;  CURRENT APPLICATION NUMBER: US/09/685,166A
;  CURRENT FILING DATE:  2000-10-10
;  NUMBER OF SEQ ID NOS: 898
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 851
;   LENGTH: 1203
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-685-166A-851

Alignment Scores:
Pred. No.:              2.21e-60      Length:          1203
Score:                  632.00        Matches:         126
Percent Similarity:     96.21%        Conservative:    1
Best Local Similarity:  95.45%        Mismatches:      1

Query Match:          93.63%      Indels:        4
DB:                   4           Gaps:          1

US-09-684-215B-4 (1-132) x US-09-685-166A-851 (1-1203)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::            ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405

RESULT 28
US-09-620-412C-348
; Sequence 348, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 348
;   LENGTH: 1464
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-348

Alignment Scores:
Pred. No.:              2.86e-60      Length:          1464
Score:                  632.00        Matches:         126
Percent Similarity:     96.21%        Conservative:    1
Best Local Similarity:  95.45%        Mismatches:      1
Query Match:            93.63%        Indels:          4
DB:                     4             Gaps:            1

US-09-684-215B-4 (1-132) x US-09-620-412C-348 (1-1464)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::            ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189
```

```
Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy         121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
               ||||||||||||||||||||||||||||||||||||
Db         370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405

RESULT 29
US-09-598-419-348
; Sequence 348, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 348
;   LENGTH: 1464
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-348

Alignment Scores:
Pred. No.:                2.86e-60        Length:         1464
Score:                    632.00          Matches:        126
Percent Similarity:       96.21%          Conservative:   1
Best Local Similarity:    95.45%          Mismatches:     1
Query Match:              93.63%          Indels:         4
DB:                       4               Gaps:           1

US-09-684-215B-4 (1-132) x US-09-598-419-348 (1-1464)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||:::                ||||||||||||
Db          82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG----------CTTCCCACCGTTCAT 129

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy         121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
               ||||||||||||||||||||||||||||||||||||
Db         370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405

RESULT 30
US-09-620-412C-332
; Sequence 332, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 332
;   LENGTH: 1557
;   TYPE: DNA
;   ORGANISM: Chlymadia trachomatis
US-09-620-412C-332

Alignment Scores:
Pred. No.:                3.1e-60         Length:         1557
Score:                    632.00          Matches:        126
Percent Similarity:       96.21%          Conservative:   1
Best Local Similarity:    95.45%          Mismatches:     1
Query Match:              93.63%          Indels:         4
DB:                       4               Gaps:           1

US-09-684-215B-4 (1-132) x US-09-620-412C-332 (1-1557)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||:::                ||||||||||||
Db          82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG----------CTTCCCACCGTTCAT 129

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy         121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
               ||||||||||||||||||||||||||||||||||||
Db         370 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 405

Search completed: April 30, 2004, 05:01:06
Job time : 1065.58 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 20:32:52 ; Search time 3670.05 Seconds
(without alignments)
1511.672 Million cell updates/sec

Title: US-09-684-215B-23
Perfect score: 651
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....SVTWQTKSGGTRTGNVTLAE 128

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061304_13185/app_query.fasta_1.1180
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_6034_@runat_29042004_061304_13185 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	651	100.0	702	6	BD251334	BD251334 Fused pro
2	651	100.0	702	6	AR403747	AR403747 Sequence
3	651	100.0	1068	6	BD274032	BD274032 Sequences
4	651	100.0	1068	6	AX005788	AX005788 Sequence
5	651	100.0	1143	6	BD274033	BD274033 Sequences
6	651	100.0	1143	6	AX005790	AX005790 Sequence
7	651	100.0	2287	6	BD251322	BD251322 Fused pro
8	651	100.0	2287	6	AR303127	AR303127 Sequence
9	651	100.0	2287	6	AR403735	AR403735 Sequence
10	651	100.0	14029	1	AE006925	AE006925 Mycobacte
11	651	100.0	341957	15	BX842572	Bx842572 Mycobacte
12	651	100.0	343050	1	BX248334	BX248334 Mycobacte
13	646	99.2	447	6	AR169152	AR169152 Sequence
14	646	99.2	447	6	AR182442	AR182442 Sequence
15	646	99.2	447	6	AR194825	AR194825 Sequence
16	646	99.2	447	6	AR233097	AR233097 Sequence
17	646	99.2	447	6	AR353302	AR353302 Sequence
18	646	99.2	447	6	AX429596	AX429596 Sequence
19	646	99.2	447	6	AX832581	AX832581 Sequence
20	646	99.2	447	6	BD006325	BD006325 Compounds
21	646	99.2	447	6	BD006445	BD006445 Compounds
22	646	99.2	447	6	BD069285	BD069285 Compounds
23	646	99.2	447	6	BD205817	BD205817 Compounds
24	646	99.2	1872	6	AR169165	AR169165 Sequence
25	646	99.2	1872	6	AR182455	AR182455 Sequence
26	646	99.2	1872	6	AR194838	AR194838 Sequence
27	646	99.2	1872	6	AR233110	AR233110 Sequence
28	646	99.2	1872	6	AR353315	AR353315 Sequence
29	646	99.2	1872	6	AX429609	AX429609 Sequence
30	646	99.2	1872	6	AX832594	AX832594 Sequence
31	646	99.2	1872	6	BD006338	BD006338 Compounds
32	646	99.2	1872	6	BD006458	BD006458 Compounds
33	646	99.2	1872	6	BD069298	BD069298 Compounds
34	646	99.2	1872	6	BD205830	BD205830 Compounds
35	608	93.4	675	6	AR261272	AR261272 Sequence
36	608	93.4	675	6	AR400535	AR400535 Sequence
37	608	93.4	675	6	AR405802	AR405802 Sequence
38	608	93.4	675	6	AX201049	AX201049 Sequence
39	608	93.4	675	6	AX267848	AX267848 Sequence
40	608	93.4	822	6	AR277645	AR277645 Sequence
41	608	93.4	822	6	AX369152	AX369152 Sequence
42	608	93.4	894	6	AX351489	AX351489 Sequence
43	608	93.4	900	6	AR220690	AR220690 Sequence
44	608	93.4	900	6	AX365960	AX365960 Sequence
45	608	93.4	915	6	AR261281	AR261281 Sequence

ALIGNMENTS

RESULT 1

```
BD251334
LOCUS       BD251334                 702 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Fused protein of Mycobacterium tuberculosis antigen and utilization
            thereof.
ACCESSION   BD251334
VERSION     BD251334.1  GI:33061104
KEYWORDS    JP 2002510494-A/13.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Skeiky,Y.A.W., Alderson,M. and Neto,A.C.
  TITLE     Fused protein of Mycobacterium tuberculosis antigen and utilization
            thereof
  JOURNAL   Patent: JP 2002510494-A 13 09-APR-2002;
            CORIXA CORP
COMMENT     OS   Artificial Sequence
            PN   JP 2002510494-A/13
            PD   09-APR-2002
            PF   07-APR-1999 JP 2000542460
            PR   07-APR-1998 US    09/056556,30-DEC-1998 US    09/223040 PI
            YASIR A W SKEIKY,MARK ALDERSON,ANTONIO CAMPOS NETO PC
            C12N15/09,A61K39/04,A61K48/00,A61P31/04,C07K14/35,C07K19/00, PC
            C12P21/02,
            PC   C12N15/00
            CC   Description of Artificial Sequence:bi-fusion
            protein Ral2-DPPD
            CC          (designated
            CC   Mtb24), reading frame 1
            CC   bi-fusion protein Ral2-DPPD (designated Mtb24),reading frame 1
            CC   reading frame 2
            CC   reading frame 3
            FH   Key             Location/Qualifiers
            FT   CDS             (1). .(696)
            FT   CDS             (2). .(700)
            FT   CDS             (3). .(701).
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Alignment Scores:
Pred. No.:                3.55e-42        Length:         702
Score:                    651.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-23 (1-128) x BD251334 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 2
AR403747
LOCUS       AR403747                 702 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 27 from patent US 6627198.
ACCESSION   AR403747
VERSION     AR403747.1  GI:40151423
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Reed,S.G., Skeiky,Y.A., Dillon,D.C., Alderson,M. and Campos-Neto,A.
  TITLE     Fusion proteins of Mycobacterium tuberculosis antigens and their
            uses
  JOURNAL   Patent: US 6627198-A 27 30-SEP-2003;
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                3.55e-42        Length:         702
Score:                    651.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-23 (1-128) x AR403747 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 3
BD274032
LOCUS       BD274032                 1068 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Sequences nuclieic acid from polypeptides exportes from
            mycobacteries, vector the comprenant and applications at diagnostic
            and the prevention from the tuberculose.
```

```
ACCESSION   BD274032
VERSION     BD274032.1  GI:33083800
KEYWORDS    JP 2002534956-A/256.
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1  (bases 1 to 1068)
  AUTHORS   Gicquel,B., Portnoi,D., Lim,E., Pelicic,V., Guigueno,A. and
            Salmoniere,Y.G.D.L.
  TITLE     Sequences nuclieic acid from polypeptides exportes from
            mycobacteries, vector the comprenant and applications at diagnostic
            and the prevention from the tuberculose
  JOURNAL   Patent: JP 2002534956-A 256 22-OCT-2002;
            INSTITUT PASTEUR
COMMENT     OS   Mycobacterium tuberculosis
            PN   JP 2002534956-A/256
            PD   22-OCT-2002
            PF   14-AUG-1998 JP 2000509849
            PR   14-AUG-1997 FR     97/10404,11-SEP-1997 FR     97/11325 PI
            BRIGITTE GICQUEL,DENIS PORTNOI,ENG-MONG LIM,VLADIMIR PELICIC, PI
            AGNES GUIGUENO,YVES GOGUET DE LA SALMONIERE
            PC   C12N15/09,A61K39/04,A61K48/00,A61P11/00,A61P31/04,C07K14/35,
            PC   C07K16/12,
            PC   C07K19/00,C12N1/21,C12P21/02,C12Q1/02,C12Q1/68,G01N33/50, PC
            G01N33/53//
            PC   (C12N15/09,C12R1:32),C12N15/00,(C12N15/00,C12R1:32) CC   50D
            FH   Key             Location/Qualifiers
            FT   CDS             (1). .(1065).
FEATURES             Location/Qualifiers
     source          1. .1068
                     /organism="Mycobacterium tuberculosis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1773"
ORIGIN
```

Alignment Scores:

Pred. No.:	5.42e-42	Length:	1068
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x BD274032 (1-1068)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 729

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 789

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 849

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 909

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 969

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1029

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||||
```

```
Db       1030 ACAGGGAACGTGACATTGGCCGAG 1053
```

```
RESULT 4
AX005788
LOCUS       AX005788               1068 bp    DNA     linear   PAT 24-AUG-2000
DEFINITION  Sequence 907 from Patent WO9909186.
ACCESSION   AX005788
VERSION     AX005788.1  GI:9928795
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1
  AUTHORS   Portnoi,D. and Guigueno,A.
  TITLE     Polypeptide nucleic sequences exported from mycobacteria, vectors
            comprising same and uses for diagnosing and preventing tuberculosis
  JOURNAL   Patent: WO 9909186-A 907 25-FEB-1999;
            PORTNOI DENIS (FR); GUIGUENO AGNES (FR)
FEATURES             Location/Qualifiers
     source          1. .1068
                     /organism="Mycobacterium tuberculosis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1773"
     CDS             1. .1068
                     /note="unnamed protein product; 50D"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC05178.1"
                     /db_xref="GI:9928796"
                     /db_xref="REMTREMBL:CAC05178"
                     /translation="MSNSRRRSLRWSWLLSVLAAVGLGLATAPAQAAPPALSQDRFAD
                     FPALPLDPSAMVAQVGPQVVNINTKLGYNNAVGAGTGIVIDPNGVVLTNNHVIAGATD
                     INAFSVGSGQTYGVDVVGYDRTQDVAVLQLRGAGGLPSAAIGGGVAVGEPVVAMGNSG
                     GQGGTPRAVPGRVVALGQTVQASDSLTGAEETLNGLIQFDAAIQPGDSGGPVVNGLGQ
                     VVGMNTAASDNFQLSQGGQGFAIPIGQAMAIAGQIRSGGGSPTVHIGPTAFLGLGVVD
                     NNGNGARVQRVVGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVT
                     WQTKSGGTRTGNVTLAEGPPA"
ORIGIN
```

Alignment Scores:

Pred. No.:	5.42e-42	Length:	1068
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x AX005788 (1-1068)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 729

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 789

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 849

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 909

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 969

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1029

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||
Db       1030 ACAGGGAACGTGACATTGGCCGAG 1053

RESULT 5
BD274033
LOCUS       BD274033               1143 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Sequences nuclieic acid from polypeptides exportes from
            mycobacteries, vector the comprenant and applications at diagnostic
            and the prevention from the tuberculose.
ACCESSION   BD274033
VERSION     BD274033.1  GI:33083801
KEYWORDS    JP 2002534956-A/257.
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1  (bases 1 to 1143)
  AUTHORS   Gicquel,B., Portnoi,D., Lim,E., Pelicic,V., Guigueno,A. and
            Salmoniere,Y.G.D.L.
  TITLE     Sequences nuclieic acid from polypeptides exportes from
            mycobacteries, vector the comprenant and applications at diagnostic
            and the prevention from the tuberculose
  JOURNAL   Patent: JP 2002534956-A 257 22-OCT-2002;
            INSTITUT PASTEUR
COMMENT     OS   Mycobacterium tuberculosis
            PN   JP 2002534956-A/257
            PD   22-OCT-2002
            PF   14-AUG-1998 JP 2000509849
            PR   14-AUG-1997 FR     97/10404,11-SEP-1997 FR     97/11325 PI
            BRIGITTE GICQUEL,DENIS PORTNOI,ENG-MONG LIM,VLADIMIR PELICIC, PI
            AGNES GUIGUENO,YVES GOGUET DE LA SALMONIERE
            PC   C12N15/09,A61K39/04,A61K48/00,A61P11/00,A61P31/04,C07K14/35,
            PC   C07K16/12,
            PC   C07K19/00,C12N1/21,C12P21/02,C12Q1/02,C12Q1/68,G01N33/50, PC
            G01N33/53//
            PC   (C12N15/09,C12R1:32),C12N15/00,(C12N15/00,C12R1:32) CC   SEQ
            ID No. 50F
            FH   Key             Location/Qualifiers
            FT   CDS             (4). .(1140).
FEATURES              Location/Qualifiers
     source           1. .1143
                      /organism="Mycobacterium tuberculosis"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:1773"
ORIGIN

Alignment Scores:
Pred. No.:              5.8e-42        Length:        1143
Score:                  651.00         Matches:       128
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     6              Gaps:          0

US-09-684-215B-23 (1-128) x BD274033 (1-1143)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 804

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 864

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 924

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 984

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1044

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1104

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||
Db       1105 ACAGGGAACGTGACATTGGCCGAG 1128

RESULT 6
AX005790
LOCUS       AX005790               1143 bp    DNA     linear   PAT 24-AUG-2000
DEFINITION  Sequence 909 from Patent WO9909186.
ACCESSION   AX005790
VERSION     AX005790.1  GI:9928797
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1
  AUTHORS   Portnoi,D. and Guigueno,A.
  TITLE     Polypeptide nucleic sequences exported from mycobacteria, vectors
            comprising same and uses for diagnosing and preventing tuberculosis
  JOURNAL   Patent: WO 9909186-A 909 25-FEB-1999;
            PORTNOI DENIS (FR); GUIGUENO AGNES (FR)
FEATURES              Location/Qualifiers
     source           1. .1143
                      /organism="Mycobacterium tuberculosis"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:1773"
     CDS              <4. .1143
                      /note="unnamed protein product; SEQ ID NO 50F"
                      /codon_start=1
                      /transl_table=11
                      /protein_id="CAC05179.1"
                      /db_xref="GI:9928798"
                      /db_xref="REMTREMBL:CAC05179"
                      /translation="KNPAARTLKAGTISDSYPDTGGYGMSNSRRRSLRWSWLLSVLAA
                      VGLGLATAPAQAAPPALSQDRFADFPALPLDPSAMVAQVGPQVVNINTKLGYNNAVGA
                      GTGIVIDPNGVVLTNNHVIAGATDINAFSVGSGQTYGVDVVGYDRTQDVAVLQLRGAG
                      GLPSAAIGGGVAVGEPVVAMGNSGGQGGTPRAVPGRVVALGQTVQASDSLTGAEETLN
                      GLIQFDAAIQPGDSGGPVVNGLGQVVGMNTAASDNFQLSQGGQGFAIPIGQAMAIAGQ
                      IRSGGGSPTVHIGPTAFLGLGVVDNNGNGARVQRVVGSAPAASLGISTGDVITAVDGA
                      PINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNVTLAEGPPA"
ORIGIN

Alignment Scores:
Pred. No.:              5.8e-42        Length:        1143
Score:                  651.00         Matches:       128
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     6              Gaps:          0

US-09-684-215B-23 (1-128) x AX005790 (1-1143)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 804

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 864

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 924

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 984

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1044

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1104

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
                ||||||||||||||||||||||||
Db       1105 ACAGGGAACGTGACATTGGCCGAG 1128

RESULT 7
BD251322
LOCUS       BD251322               2287 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Fused protein of Mycobacterium tuberculosis antigen and utilization
            thereof.
ACCESSION   BD251322
VERSION     BD251322.1  GI:33061092
KEYWORDS    JP 2002510494-A/1.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 2287)
  AUTHORS   Skeiky,Y.A.W., Alderson,M. and Neto,A.C.
  TITLE     Fused protein of Mycobacterium tuberculosis antigen and utilization
            thereof
  JOURNAL   Patent: JP 2002510494-A 1 09-APR-2002;
            CORIXA CORP
COMMENT     OS   Artificial Sequence
            PN   JP 2002510494-A/1
            PD   09-APR-2002
            PF   07-APR-1999 JP 2000542460
            PR   07-APR-1998 US    09/056556,30-DEC-1998 US    09/223040 PI
            YASIR A W SKEIKY,MARK ALDERSON,ANTONIO CAMPOS NETO PC
            C12N15/09,A61K39/04,A61K48/00,A61P31/04,C07K14/35,C07K19/00, PC
            C12P21/02,
            PC   C12N15/00
            CC   Description of Artificial Sequence:tri-fusion protein Ra12- CC
                 TbH9-Ra35
            CC   (designated Mtb32A)
            CC   n = g, a, c or t
            CC   n = g, a, c or t
            CC   n = g, a, c or t
            FH   Key             Location/Qualifiers
            FT   modified_base   (30)
            FT   modified_base   (33)
            FT   CDS             (42). .(2231)
            FT   modified_base   (2270).
FEATURES             Location/Qualifiers
     source          1. .2287
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Alignment Scores:
Pred. No.:              1.16e-41        Length:         2287
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-23 (1-128) x BD251322 (1-2287)
```

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
                ||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAG 446

RESULT 8
AR303127
LOCUS       AR303127               2287 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 1 from patent US 6544522.
ACCESSION   AR303127
VERSION     AR303127.1  GI:31691855
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2287)
  AUTHORS   Skeiky,Y., Alderson,M. and Campos-Neto,A.
  TITLE     Fusion proteins of mycobacterium tuberculosis antigens and their
            uses
  JOURNAL   Patent: US 6544522-A 1 08-APR-2003;
FEATURES             Location/Qualifiers
     source          1. .2287
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              1.16e-41        Length:         2287
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-23 (1-128) x AR303127 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAG 446

RESULT 9
AR403735
LOCUS       AR403735                2287 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 1 from patent US 6627198.
ACCESSION   AR403735
VERSION     AR403735.1  GI:40151411
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2287)
  AUTHORS   Reed,S.G., Skeiky,Y.A., Dillon,D.C., Alderson,M. and Campos-Neto,A.
  TITLE     Fusion proteins of Mycobacterium tuberculosis antigens and their
            uses
  JOURNAL   Patent: US 6627198-A 1 30-SEP-2003;
FEATURES             Location/Qualifiers
     source          1. .2287
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                1.16e-41        Length:          2287
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-23 (1-128) x AR403735 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAG 446

RESULT 10
AE006925
LOCUS       AE006925                14029 bp    DNA     linear   BCT 27-APR-2001
DEFINITION  Mycobacterium tuberculosis CDC1551,  section 11 of 280 of the
            complete genome.
ACCESSION   AE006925 AE000516
VERSION     AE006925.1  GI:13879610
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis CDC1551
  ORGANISM  Mycobacterium tuberculosis CDC1551
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1  (bases 1 to 14029)
  AUTHORS   Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
            Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
            Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
            Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
            Gill,J., Mikula,A. and Bishai,W.
  TITLE     Whole genome comparison of Mycobacterium tuberculosis clinical and
            laboratory strains
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 14029)
  AUTHORS   Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
            Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
            Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
            Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
            Gill,J., Mikula,A. and Bishai,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2001) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .14029
                     /organism="Mycobacterium tuberculosis CDC1551"
                     /mol_type="genomic DNA"
                     /strain="CDC1551"
                     /db_xref="taxon:83331"
                     /note="clinical strain"
     gene            complement(99. .533)
                     /gene="MT0129"
     CDS             complement(99. .533)
                     /gene="MT0129"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAK44353.1"
                     /db_xref="GI:13879611"
                     /translation="MGEFDPKLRFAQSPVARLATSTPDGTPHLVPVVFALGARRPAEA
                     TGADVIYTAVDAKRKTTQRLRRLANLEHNPRASVLVDSYADDWTQLWWVRADGVAAIH
                     RDGEVMRAAYRLLRAKYTQYQSVPLNGPVIAIAVQRWASRHA"
     gene            complement(643. .1062)
                     /gene="MT0130"
     CDS             complement(643. .1062)
                     /gene="MT0130"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAK44354.1"
                     /db_xref="GI:13879612"
                     /translation="MASSSSLPLLRHCVSDDQVTVVGFDGDDLGKTARRIAALVVQRA
                     IFLNDRNTAVAHSGDDAVLGHAVLPGVPRDPDPLHASSMYSILGMCQSVNGRPFDAIA
                     LVSVRLCHVQTDPTDSCGGRDRPGQLPCAPLDYHRHH"
     gene            1047. .1415
                     /gene="MT0131"
     CDS             1047. .1415
                     /gene="MT0131"
                     /note="identified by match to  protein family HMM"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA-binding protein, CopG family"
                     /protein_id="AAK44355.1"
```

/db_xref="GI:13879613"
/translation="MTKKPRNPADYVIGDDVEVSDVDLKQEEVYVDGERLTDERVEQM
ASESLRLAREREANLIPGGKSLSGGSAHSPAVQVVVSKATHAKLKELARSRKMSVSKL
LRPVLDEFVQRETGRILPRR"
gene 1682. .3367
/gene="MT0132"
CDS 1682. .3367
/gene="MT0132"
/note="similar to GB:AL123456; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="PE_PGRS family protein"
/protein_id="AAK44356.1"
/db_xref="GI:13879614"
/translation="MFAGGGAGGLGRCVMSFVSVAPEIVVAAATDLAGIGSAISAANA
AAAAPTTAVLAAGADEVSAAIAALFSGHAQAYQALSAQAAAFHQQFVQTLAGGAGAYA
AAEAQVEQQLLAAINAPTQALLGRPLIGNGADGAPGTGQAGGAGGILYGNGGNGGSGA
AGQAGGAGGPAGLIGHGGSGGAGGSGAAGGAGGHGGWLWGNGGVGGSGGAGVGAGVAG
GHGGAGGAAGLWGAGGGGGNGGNGADANIVSGGDGGLGGAGGGGGWLYGDGGAGGHGG
QGAGGGAGGAGGDGGQGGAGRGLWGTGGAGGHGGQGGGTGGPPLPGQAGMGAAGGAGG
LIGNGGAGGDGGVGASGGVAGVXGAGGNAMLIGHGGAGGAGGDSSFANGAAGGAGGAG
GHLFGNGGSGGHGGAVTAGNTGIGGAGGVGGDARLIGHGGAGGAGGDRAGALVGRDGG
PGGNGGAGGQLYGNGGDGGPGGQGGQAFGANNIGGTGGAGGNGGPAILSGNGGNGGAG
GAGGAGGAGGGAGGVGGAGGAPGTGGTLQAAVSGLVTALFGAPGQPGDTGQPG"
gene 3519. .4586
/gene="MT0133"
CDS 3519. .4586
/gene="MT0133"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="serine protease, putative"
/protein_id="AAK44357.1"
/db_xref="GI:13879615"
/translation="MSNSRRRSLRWSWLLSVLAAVGLGLATAPAQAAPPALSQDRFAD
FPALPLDPSAMVAQVGPQVVNINTKLGYNNAVGAGTGIVIDPNGVVLTNNHVIAGATD
INAFSVGSGQTYGVDVVGYDRTQDVAVLQLRGAGGLPSAAIGGGVAVGEPVVAMGNSG
GQGGTPRAVPGRVVALGQTVQASDSLTGAEETLNGLIQFDAAIQPGDSGGPVVNGLGQ
VVGMNTAASDNFQLSQGGQGFAIPIGQAMAIAGQIRSGGGSPTVHIGPTAFLGLGVVD
NNGNGARVQRVVGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVT
WQTKSGGTRTGNVTLAEGPPA"
gene 4695. .6500
/gene="MT0134"
CDS 4695. .6500
/gene="MT0134"
/note="similar to GP:2808807; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="alpha-amylase family protein"
/protein_id="AAK44358.1"
/db_xref="GI:13879616"
/translation="MNEAEHSVEHPPVQGSHVEGGVVEHPDAKDFGSAAALPADPTWF
KHAVFYEVLVRAFFDASADGSGDLRGLIDRLDYLQWLGIDCIWLPPFYDSPLRDGGYD
IRDFYKVLPEFGTVDDFVALVDAAHRRGIRIITDLVMNHTSESHPWFQESRRDPDGPY
GDYYVWSDTSERYTDARIIFVDTEESNWSFDPVRRQFYWHRFFSHQPDLNYDNPAVQE
AMIDVIRFWLGLGIDGFRLDAVPYLFEREGTNCENLPETHAFLKRVRKVVDDEFPGRV
LLAEANQWPGDVVEYFGDPNTGGDECHMAFHFPLMPRIFMAVRRESRFPISEIIAQTP
PIPDMAQWGIFLRNHDELTLEMVTDEERDYMYAEYAKDPRMKANVGIRRRLAPLLDND
RNQIELFTALLLSLPGSPVLYYGDEIGMGDVIWLGDRDGVRIPMQWTPDRNAGFSTAN
PGRLYLPPSQDPVYGYQAVNVEAQRDTSTSLLNFTRTMLAVRRRHPAFAVGAFQELGG
SNPSVLAYVRQVAGDDGDTVLCVNNLSRFPQPIELDLQQWTNYTPVELTGHVEFPRIG
QVPYLLTLPGHGFYWFQLTTHEVGAPPTCGGERRL"
gene 6603. .7970
/gene="MT0135"
CDS 6603. .7970
/gene="MT0135"
/note="similar to GP:2808801; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="pep2 protein"
/protein_id="AAK44359.1"
/db_xref="GI:13879617"
/translation="MTRSDTLATKLPWSDWLPRQRWYAGRNRELATVKPGVVVALRHN
LDLVLVDVTYTDGATERYQVLVGWDFEPASEYGTKAAIGVADDRTGFDALYDVAGPQF
LLSLIVSSAVCGTSTGEVTFTREPDVELPFAAQPRVCDAEQSNTSVIFDRRAILKVFR
RVSSGINPDIELNRVLTRAGNPHVARLLGAYQFGRPNRSPTDALAYALGMVTEYEANA
AEGWAMATASVRDLFAEGDLYAHEVGGDFAGESYRLGEAVASVHATLADSLGTAQATF
PVDRMLARLSSTVAVVPELREYAPTIEQQFQKLAAEAITVQRVHGDLHLGQVLRTPES
WLLIDFEGEPGQPLDERRAPDSPLRDVAGVLRSFEYAAYGPLVDQATDKQLAARAREW
VERNRAAFCDGYAVASGIDPRDSALLLGAYELDKAVYETGYETRHRPGWLPIPLRSIA
RLTAS"
gene 8038. .8817
/gene="MT0136"
CDS 8038. .8817
/gene="MT0136"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44360.1"
/db_xref="GI:13879618"
/translation="MQREIYDGEARLSWVLAALAGILGATAFTHSAGYFVTFMTGNSQ
RAVLGLFGDDAWMSVTASLLILFFVAGVVIASVCRRHFWAAHPHGPTVLTTFSLIFAA
GVDIMLGGWHESMLDFVPILFVVFGIGALNTSFVKDGEVSVPLSYVTGTLVKMGQGIE
RHLAGGKVEDWLGYFLLHASFVLGAAAGGAISMVVTGPQMLAVAAVVCAATTGYTYLH
ADRRGLVNQKRPQPGKRLFRALRRGELDSGTSTPATNYGSS"
gene complement(8949. .9989)
/gene="MT0137"
CDS complement(8949. .9989)
/gene="MT0137"
/note="similar to SP:P31953 GB:X57229 PID:48828
PID:1567734 PID:1877254; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="esterase, putative, antigen 85-C"
/protein_id="AAK44361.1"
/db_xref="GI:13879619"
/translation="MGLVVAMTFFEQVRRLRSAATTLPRRLAIAAMGAVLVYGLVGTF
GGPATAGAFSRPGLPVEYLQVPSASMGRDIKVQFQGGGPHAVYLLDGLRAQDDYNGWD
INTPAFEEYYQSGLSVIMPVGGQSSFYTDWYQPSQSNGQNYTYKWETFLTREMPAWLQ
ANKGVSPTGNAAVGLSMSGGSALILAAYYPQQFPYAASLSGFLNPSEGWWPTLIGLAM
NDSGGYNANSMWGPSSDPAWKRNDPMVQIPRLVANNTRIWVYCGNGTPSDLGGDNIPA
KFLEGLTLRTNQTFRDTYAADGGRNGVFNFPPNGTHSWPYWNEQLVAMKADIQHVLNG
ATPPAAPAAPAAPAA"
gene 10218. .10673
/gene="MT0138"
CDS 10218. .10673
/gene="MT0138"
/note="similar to SP:P08634 PID:46222; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="nodulation protein N-related protein"
/protein_id="AAK44362.1"
/db_xref="GI:13879620"

Alignment Scores:

Pred. No.:	7.2e-41	Length:	14029
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-684-215B-23 (1-128) x AE006925 (1-14029)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
Db 4188 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 4247

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40

```
Db       4248 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 4307

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4308 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 4367

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4368 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 4427

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4428 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 4487

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4488 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 4547

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db       4548 ACAGGGAACGTGACATTGGCCGAG 4571

RESULT 11
BX842572
ID   BX842572   standard; circular genomic DNA; PRO; 341957 BP.
XX
AC   BX842572; AL021427; AL021428; AL021926; AL021927; AL021928; AL021929;
AC   AL021930; AL123456; Z74410; Z80233; Z80775; Z86089; Z92669; Z92770; Z96071;
AC   Z97050;
XX
SV   BX842572.1
XX
DT   21-NOV-2003 (Rel. 77, Created)
DT   21-NOV-2003 (Rel. 77, Last updated, Version 1)
XX
DE   Mycobacterium tuberculosis H37Rv complete genome; segment 1/13
XX
KW   complete genome.
XX
OS   Mycobacterium tuberculosis H37Rv
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium tuberculosis complex; Mycobacterium tuberculosis.
XX
RN   [1]
RX   MEDLINE; 98295987.
RX   PUBMED; 9634230.
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry III C.E., Tekaia F., Badcock K.,
RA   Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K.,
RA   Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K.,
RA   Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J.,
RA   Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S.,
RA   Squares S., Sqares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the complete
RT   genome sequence";
RL   Nature 393:537-544(1998).
XX
RN   [2]
RX   PUBMED; 12368430.
RA   Camus J.C., Pryor M.J., Medigue C., Cole S.T.;
RT   "Re-annotation of the genome sequence of Mycobacterium tuberculosis H37Rv";
RL   Microbiology 148:2967-2973(2002).
XX
RN   [3]
RP   1-341957
RA   Parkhill J.;
RT   ;
RL   Submitted (11-JUN-1998) to the EMBL/GenBank/DDBJ databases.
RL   Submitted on behalf of the Mycobacterium tuberculosis sequencing and
RL   mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton,
RL   Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut
RL   Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail:
RL   parkhill@sanger.ac.uk
XX
CC   Notes:
CC   Details of M. tuberculosis sequencing at the Sanger Centre
CC   are available on the World Wide Web.
CC   (URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/)
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .341957
FT                   /db_xref="taxon:83332"
FT                   /mol_type="genomic DNA"
FT                   /organism="Mycobacterium tuberculosis H37Rv"
FT                   /strain="H37Rv"
FT   CDS             1. .1524
FT                   /evidence=EXPERIMENTAL
FT                   /note="Rv0001, (MT0001, MTV029.01, P49993), len: 507 aa.
FT                   dnaA, chromosomal replication initiator protein (see
FT                   citations below), equivalent to other Mycobacterial
FT                   CHROMOSOMAL REPLICATION INITIATOR PROTEINS e.g.
FT                   P46388|DNAA_MYCLE from Mycobacterium leprae (502 aa);
FT                   Q9L7L7|DNAA_MYCPA from Mycobacterium paratuberculosis (509
FT                   aa); P49990|DNAA_MYCAV from Mycobacterium avium (508 aa);
FT                   P49992|DNAA_MYCSM from Mycobacterium smegmatis (504 aa);
FT                   etc. Also highly similar to others except in N-terminus
FT                   e.g. Q9ZH75|DNAA_STRCH CHROMOSOMAL REPLICATION INITIATOR
FT                   PROTEIN from Streptomyces chrysomallus (624 aa);
FT                   Q9ZH76|DNAA_STRRE from Streptomyces reticuli (643 aa);
FT                   DNAA_ECOLI|P03004|B3702 chromosomal replication initiator
FT                   protein from Escherichia coli strain K12 (467 aa), FASTA
FT                   scores: opt: 986, E(): 0, (43.2% identity in 389 aa
FT                   overlap); etc. Contains PS00017 ATP/GTP-binding site motif
FT                   A (P-loop) and PS01008 DnaA protein signature. BELONGS TO
FT                   THE DNAA FAMILY. Note that the first base of this gene has
FT                   been taken as base 1 of the Mycobacterium tuberculosis
FT                   H37Rv genomic sequence."
FT                   /transl_table=11
FT                   /function="PLAYS AN IMPORTANT ROLE IN THE INITIATION AND
FT                   REGULATION OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN
FT                   OF REPLICATION; IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA
FT                   AT A 9 BP CONSENSUS (DNAA BOX): 5'-TTATC(C/A)A(C/A)A-3'.
FT                   DNAA BINDS TO ATP AND TO ACIDIC PHOSPHOLIPIDS. DNAA PROTEIN
FT                   BINDS THE ORIGIN OF REPLICATION (oriC), ATP AND ADP, AND
FT                   EXHIBITED WEAK ATPase ACTIVITY."
FT                   /gene="dnaA"
FT                   /locus_tag="Rv0001"
FT                   /product="CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA"
FT                   /protein_id="CAA16238.1"
FT                   /translation="MTDDPGSGFTTVWNAVVSELNGDPKVDDGPSSDANLSAPLTPQQR
FT                   AWLNLVQPLTIVEGFALLSVPSSFVQNEIERHLRAPITDALSRRLGHQIQLGVRIAPPA
FT                   TDEADDTTVPPSENPATTSPDTTTDNDEIDDSAAARGDNQHSWPSYFTERPHNTDSATA
FT                   GVTSLNRRYTFDTFVIGASNRFAHAAALAIAEAPARAYNPLFIWGESGLGKTHLLHAAG
FT                   NYAQRLFPGMRVKYVSTEEFTNDFINSLRDDRKVAFKRSYRDVDVLLVDDIQFIEGKEG
FT                   IQEEFFHTFNTLHNANKQIVISSDRPPKQLATLEDRLRTRFEWGLITDVQPPELETRIA
FT                   ILRKKAQMERLAVPDDVLELIASSIERNIRELEGALIRVTAFASLNKTPIDKALAEIVL
FT                   RDLIADANTMQISAATIMAATAEYFDTTVEELRGPGKTRALAQSRQIAMYLCRELTDLS
FT                   LPKIGQAFGRDHTTVMYAQRKILSEMAERREVFDHVKELTTRIRQRSKR"
FT   misc_feature    622. .645
FT                   /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
FT   misc_feature    1384. .1440
FT                   /note="PS01008 DnaA protein signature"
FT   CDS             2052. .3260
FT                   /evidence=EXPERIMENTAL
FT                   /note="Rv0002, (MTV029.02, MTCY10H4.0), len: 402 aa. dnaN,
FT                   DNA polymerase III (beta chain) (EC 2.7.7.7) (see citations
FT                   below), equivalent to other Mycobacterial DNA POLYMERASES
FT                   III BETA CHAIN e.g. NP_301130.1|NC_002677 from
FT                   Mycobacterium leprae (399 aa); Q9L7L6|DP3B_MYCPA from
FT                   Mycobacterium avium subsp. paratuberculosis (399 aa);
FT                   P52851|DP3B_MYCSM from Mycobacterium smegmatis (397 aa);
FT                   etc. Also highly similar to others e.g. P27903|DP3B_STRCO
FT                   DNA POLYMERASE III BETA CHAIN from Streptomyces coelicolor
FT                   (376 aa), FASTA scores: opt: 1189, E(): 0, (52.8% identity
```

```
FT                   in 337 aa overlap); P21174|DP3B_MICLU from Micrococcus
FT                   luteus (310 aa); P52023|DP3B_SYNP7 from Synechococcus sp.
FT                   strain PCC 7942 (375 aa); etc. Overlaps and extends CDS in
FT                   neighbouring cosmid MTCY10H4.01."
FT                   /transl_table=11
FT                   /EC_number="2.7.7.7"
FT                   /function="DNA POLYMERASE III IS A COMPLEX, MULTICHAIN
FT                   ENZYME RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN
FT                   BACTERIA. THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5'
FT                   EXONUCLEASE ACTIVITY. THE BETA CHAIN IS REQUIRED FOR
FT                   INITIATION OF REPLICATION ONCE IT IS CLAMPED ONTO DNA, IT
FT                   SLIDES FREELY (BIDIRECTIONAL AND ATP-INDEPENDENT) ALONG
FT                   DUPLEX DNA [CATALYTIC ACTIVITY: N deoxynucleoside
FT                   triphosphate = N diphosphate + {DNA}N]."
FT                   /gene="dnaN"
FT                   /locus_tag="Rv0002"
FT                   /product="DNA POLYMERASE III (BETA CHAIN) DNAN (DNA
FT                   NUCLEOTIDYLTRANSFERASE)"
FT                   /protein_id="CAA16239.1"
FT                   /translation="MDAATTRVGLTDLTFRLLRESFADAVSWVAKNLPARPAVPVLSGV
FT                   LLTGSDNGLTISGFDYEVSAEAQVGAEIVSPGSVLVSGRLLSDITRALPNKPVDVHVEG
FT                   NRVALTCGNARFSLPTMPVEDYPTLPTLPEETGLLPAELFAEAISQVAIAAGRDDTLPM
FT                   LTGIRVEILGETVVLAATDRFRLAVRELKWSASSPDIEAAVLVPAKTLAEAAKAGIGGS
FT                   DVRLSLGTGPGVGKDGLLGISGNGKRSTTRLLDAEFPKFRQLLPTEHTAVATMDVAELI
FT                   EAIKLVALVADRGAQVRMEFADGSVRLSAGADDVGRAEEDLVVDYAGEPLTIAFNPTYL
FT                   TDGLSSLRSERVSFGFTTAGKPALLRPVSGDDRPVAGLNGNGPFPAVSTDYVYLLMPVR
FT                   LPG"
FT   CDS             3280. .4437
FT                   /evidence=EXPERIMENTAL
FT                   /note="Rv0003, (MTCY10H4.01), len: 385 aa. recF, DNA
FT                   replication and repair protein (see citations below),
FT                   equivalent to others Mycobacterial DNA replication and
FT                   repair proteins e.g. NP_301131.1|NC_002677 from
FT                   Mycobacterium leprae (385 aa); Q9L7L5|RECF_MYCPA from
FT                   Mycobacterium avium subsp. paratuberculosis (385 aa);
FT                   P50916|RECF_MYCSM from Mycobacterium smegmatis (384 aa);
FT                   etc. Also highly similar to others e.g. P36176|RECF_STRCO
FT                   DNA REPLICATION AND REPAIR PROTEIN from Streptomyces
FT                   coelicolor (373 aa); NP_440892.1|NC_000911 from
FT                   Synechocystis sp. strain PCC 6803 (384 aa);
FT                   NP_469352.1|NC_003212 from Listeria innocua (370 aa); etc.
FT                   Contains PS00017 ATP/GTP-binding site motif A (P-loop),
FT                   PS00617 RecF protein signature 1, and PS00618 RecF protein
FT                   signature 2. BELONGS TO THE RECF FAMILY."
FT                   /transl_table=11
FT                   /function="THE RECF PROTEIN IS INVOLVED IN DNA METABOLISM
FT                   AND RECOMBINATION; IT IS REQUIRED FOR DNA REPLICATION AND
FT                   NORMAL SOS INDUCIBILITY. RECF BINDS PREFERENTIALLY TO
FT                   SINGLE-STRANDED, LINEAR DNA. IT ALSO SEEMS TO BIND ATP."
FT                   /gene="recF"
FT                   /locus_tag="Rv0003"
FT                   /product="DNA REPLICATION AND REPAIR PROTEIN RECF
FT                   (SINGLE-STRAND DNA BINDING PROTEIN)"
FT                   /protein_id="CAB02424.1"
FT                   /translation="MYVRHLGLRDFRSWACVDLELHPGRTVFVGPNGYGKTNLIEALWY
FT                   STTLGSHRVSADLPLIRVGTDRAVISTIVVNDGRECAVDLEIATGRVNKARLNRSSVRS
FT                   TRDVVGVLRAVLFAPEDLGLVRGDPADRRRYLDDLAIVRRPAIAAVRAEYERVLRQRTA
FT                   LLKSVPGARYRGDRGVFDTLEVWDSRLAEHGAELVAARIDLVNQLAPEVKKAYQLLAPE
FT                   SRSASIGYRASMDVTGPSEQSDIDRQLLAARLLAALAARRDAELERGVCLVGPHRDDLI
FT                   LRLGDQPAKGFASHGEAWSLAVALRLAAYQLLRVDGGEPVLLLDDVFAELDVMRRRALA
FT                   TAAESAEQVLVTAAVLEDIPAGWDARRVHIDVRADDTGSMSVVLP"
FT   misc_feature    3367. .3390
FT                   /note="PS00017 ATP/GTP-binding site motif A"
FT   misc_feature    3634. .3690
FT                   /note="PS00617 RecF protein signature 1"
FT   misc_feature    4243. .4296
FT                   /note="PS00618 RecF protein signature 2"
FT   CDS             4434. .4997
FT                   /evidence=EXPERIMENTAL
FT                   /note="Rv0004, (MTCY10H4.02), len: 187 aa. Conserved
FT                   hypothetical protein (see Salazar et al., 1996), highly
FT                   similar, but longer 21 aa in N-terminus, to
FT                   AAF33696.1|AF222789 unknown protein from Mycobacterium
FT                   avium subsp. paratuberculosis (166 aa); and highly similar
FT                   to NP_301132.1|NC_002677 conserved hypothetical protein
FT                   from Mycobacterium leprae (189 aa); S70990 hypothetical
FT                   protein from Mycobacterium smegmatis (194 aa). Also highly
FT                   similar, except in N-terminal part, to
```

```
Alignment Scores:
Pred. No.:                1.78e-39        Length:          341957
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       15              Gaps:            0

US-09-684-215B-23 (1-128) x BX842572 (1-341957)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151817 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 151876

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151877 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 151936

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151937 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 151996

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151997 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 152056

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152057 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 152116

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152117 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 152176

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db     152177 ACAGGGAACGTGACATTGGCCGAG 152200

RESULT 12
BX248334
LOCUS       BX248334               343050 bp    DNA     linear   BCT 11-JUN-2003
DEFINITION  Mycobacterium bovis subsp. bovis AF2122/97 complete genome; segment
            1/14.
ACCESSION   BX248334 BX248333
VERSION     BX248334.1  GI:31616762
KEYWORDS    complete genome.
SOURCE      Mycobacterium bovis subsp. bovis AF2122/97
  ORGANISM  Mycobacterium bovis subsp. bovis AF2122/97
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1
  AUTHORS   Garnier,T., Eiglmeier,K., Camus,J.-C., Medina,N., Mansoor,H.,
            Pryor,M., Duthoy,S., Grondin,S., Lacroix,C., Monsempe,C., Simon,S.,
            Harris,B., Atkin,R., Doggett,J., Mayes,R., Keating,L.,
            Wheeler,P.R., Parkhill,J., Barrell,B.G., Cole,S.T., Gordon,S.V. and
            Hewinson,G.
  TITLE     The complete genome sequence of Mycobacterium bovis
  JOURNAL   Online Publication
  REMARK    PNAS 10.1073/pnas.1130426100 { Microbiology }
REFERENCE   2  (bases 1 to 343050)
  AUTHORS   Garnier,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-2003) Garnier T., Unite de Genetique Moleculaire
            Bacterienne Institut Pasteur 28,rue du Dr Roux 75724 PARIS cedex
            15, France. e-mail:tgarnier@pasteur.fr Submitted on behalf of the
            Mycobacterium bovis sequencing teams, TB Research Group, Veterinary
```

Laboratories Agency Weybridge, Woodham Lane, New Haw,Addlestone, Surrey KT15 3NB, UK. Sanger Centre, Wellcome Trust Genome Campus, Hinxton,Cambridge CB10 1SA, UK. PT4 Annotation, Genopole, Institut Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France. Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

FEATURES Location/Qualifiers

source 1. .343050
/organism="Mycobacterium bovis subsp. bovis AF2122/97"
/mol_type="genomic DNA"
/strain="AF2122/97"
/db_xref="taxon:233413"

gene 1. .1524
/gene="dnaA"
/locus_tag="Mb0001"

CDS 1. .1524
/gene="dnaA"
/locus_tag="Mb0001"
/note="Mb0001, dnaA, len: 507 aa. Equivalent to Rv0001, len: 507 aa, from Mycobacterium tuberculosis strain H37Rv, (99.6% identity in 507 aa overlap). dnaA, chromosomal replication initiator protein (see citations below), equivalent to other Mycobacterial CHROMOSOMAL REPLICATION INITIATOR PROTEINS e.g. P46388|DNAA_MYCLE from Mycobacterium leprae (502 aa); Q9L7L7|DNAA_MYCPA from Mycobacterium paratuberculosis (509 aa); P49990|DNAA_MYCAV from Mycobacterium avium (508 aa); P49992|DNAA_MYCSM from Mycobacterium smegmatis (504 aa); etc. Also highly similar to others except in N-terminus e.g. Q9ZH75|DNAA_STRCH CHROMOSOMAL REPLICATION INITIATOR PROTEIN from Streptomyces chrysomallus (624 aa); Q9ZH76|DNAA_STRRE from Streptomyces reticuli (643 aa); DNAA_ECOLI|P03004|B3702 chromosomal replication initiator protein from Escherichia coli strain K12 (467 aa), FASTA scores: opt: 986, E(): 0, (43.2% identity in 389 aa overlap); etc. Contains PS00017 ATP/GTP-binding site motif A (P-loop) and PS01008 DnaA protein signature. BELONGS TO THE DNAA FAMILY. Note that the first base of this gene has been taken as base 1 of the Mycobacterium bovis genomic sequence."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA"
/protein_id="CAD92863.1"
/db_xref="GI:31616763"
/translation="MTDDPGSGFTTVWNAVVSELNGDPKVDDGPSSDANLSAPLTPQQ
RAWLNLVQPLTIVEGFALLSVPSSFVQNEIERHLRAPITDALSRRLGHQIQLGVRIAP
PATDEADDTTVPPSENPATTSPDTTTDNDEIDDSAAARGDNQHSWPSYFTERPRNTDS
ATAGVTSLNRRYTFDTFVIGASNRFAHAAALAIAEAPARAYNPLFIWGESGLGKTHLL
HAAGNYAQRLFPGMRVKYVSTEEFTNDFINSLRDDRKVAFKRSYRDVDVLLVDDIQFI
EGKEGIQEEFFHTFNTLHNANKQIVISSDRPPKQLATLEDRLRTRFEWGLITDVQPPE
LETRIAILRKKAQMERLAIPDDVLELIASSIERNIRELEGALIRVTAFASLNKTPIDK
ALAEIVLRDLIADANTMQISAATIMAATAEYFDTTVEELRGPGKTRALAQSRQIAMYL
CRELTDLSLPKIGQAFGRDHTTVMYAQRKILSEMAERREVFDHVKELTTRIRQRSKR"

gene 2052. .3260
/gene="dnaN"
/locus_tag="Mb0002"

CDS 2052. .3260
/gene="dnaN"
/locus_tag="Mb0002"
/EC_number="2.7.7.7"
/note="Mb0002, dnaN, len: 402 aa. Equivalent to Rv0002, len: 402 aa, from Mycobacterium tuberculosis strain H37Rv, (99.8% identity in 402 aa overlap). dnaN, DNA polymerase III (beta chain) (EC 2.7.7.7) (see citations below), equivalent to other Mycobacterial DNA POLYMERASES III BETA CHAIN e.g. NP_301130.1|NC_002677 from Mycobacterium leprae (399 aa); Q9L7L6|DP3B_MYCPA from Mycobacterium avium subsp. paratuberculosis (399 aa); P52851|DP3B_MYCSM from Mycobacterium smegmatis (397 aa); etc. Also highly similar to others e.g. P27903|DP3B_STRCO DNA POLYMERASE III BETA CHAIN from Streptomyces coelicolor (376 aa), FASTA scores: opt: 1189, E(): 0, (52.8% identity in 337 aa overlap); P21174|DP3B_MICLU from Micrococcus luteus (310 aa); P52023|DP3B_SYNP7 from Synechococcus sp. strain PCC 7942 (375 aa); etc. Overlaps and extends CDS in neighbouring cosmid MTCY10H4.01."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="DNA POLYMERASE III (BETA CHAIN) DNAN (DNA NUCLEOTIDYLTRANSFERASE)"
/protein_id="CAD92864.1"
/db_xref="GI:31616764"
/translation="MDAATTRVGLTDLTFRLLRESFADAVSWVAKNLPARPAVPVLSG
VLLTGSDNGLTISGFDYEVSAEAQVGAEIVSPGSVLVSGRLLSDITRALPNKPVGVHV
EGNRVALTCGNARFSLPTMPVEDYPTLPTLPEETGLLPAELFAEAISQVAIAAGRDDT
LPMLTGIRVEILGETVVLAATDRFRLAVRELKWSASSPDIEAAVLVPAKTLAEAAKAG
IGGSDVRLSLGTGPGVGKDGLLGISGNGKRSTTRLLDAEFPKFRQLLPTEHTAVATMD
VAELIEAIKLVALVADRGAQVRMEFADGSVRLSAGADDVGRAEEDLVVDYAGEPLTIA
FNPTYLTDGLSSLRSERVSFGFTTAGKPALLRPVSGDDRPVAGLNGNGPFPAVSTDYV
YLLMPVRLPG"

gene 3280. .4437
/gene="recF"
/locus_tag="Mb0003"

CDS 3280. .4437
/gene="recF"
/locus_tag="Mb0003"
/note="Mb0003, recF, len: 385 aa. Equivalent to Rv0003, len: 385 aa, from Mycobacterium tuberculosis strain H37Rv, (99.5% identity in 385 aa overlap). recF, DNA replication and repair protein (see citations below), equivalent to others Mycobacterial DNA replication and repair proteins e.g. NP_301131.1|NC_002677 from Mycobacterium leprae (385 aa); Q9L7L5|RECF_MYCPA from Mycobacterium avium subsp. paratuberculosis (385 aa); P50916|RECF_MYCSM from Mycobacterium smegmatis (384 aa); etc. Also highly similar to others e.g. P36176|RECF_STRCO DNA REPLICATION AND REPAIR PROTEIN from Streptomyces coelicolor (373 aa); NP_440892.1|NC_000911 from Synechocystis sp. strain PCC 6803 (384 aa); NP_469352.1|NC_003212 from Listeria innocua (370 aa); etc. Contains PS00017 ATP/GTP-binding site motif A (P-loop), PS00617 RecF protein signature 1, and PS00618 RecF protein signature 2. BELONGS TO THE RECF FAMILY."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="DNA REPLICATION AND REPAIR PROTEIN RECF (SINGLE-STRAND DNA BINDING PROTEIN)"
/protein_id="CAD92865.1"
/db_xref="GI:31616765"
/translation="MYVRHLGLRDFRSWACVDLELHPGRTVFVGPNGYGKTNLIEALW
YSTTLGSHRVSADLPLIRVGTDRAVISTIVVNDGRECAVDLEIATGRVNKARLNRSSV
RSTRDVVGVLRAVLFAPEDLGLVRGDPADRRRYLDDLAIVRRPAIAAVRAEYERVVRQ
RTALLKSVPGARYRGDRGVFDTLEVWDSRLAEHGAELVAARIDLVNQLAPEVKKAYQL
LAPESRSASIGYRASMDVTGPSEQSDTDRQLLAARLLAALAARRDAELERGVCLVGPH
RDDLILRLGDQPAKGFASHGEAWSLAVALRLAAYQLLRVDGGEPVLLLDDVFAELDVM
RRRALATAAESAEQVLVTAAVLEDIPAGWDARRVHIDVRADDTGSMSVVLP"

gene 4434. .4997
/locus_tag="Mb0004"

CDS 4434. .4997
/locus_tag="Mb0004"
/note="Mb0004, -, len: 187 aa. Equivalent to Rv0004, len: 187 aa, from Mycobacterium tuberculosis strain H37Rv, (99.5% identity in 187 aa overlap). Conserved hypothetical protein (see citation below), highly similar, but longer 21 aa in N-terminus, to AAF33696.1|AF222789 unknown protein from Mycobacterium avium subsp. paratuberculosis (166 aa); and highly similar to NP_301132.1|NC_002677 conserved hypothetical protein from Mycobacterium leprae (189 aa); S70990 hypothetical protein from Mycobacterium smegmatis (194 aa). Also similar to in C-terminus to C-terminal part of P35925|YREG_STRCO HYPOTHETICAL 19.8 KDA PROTEIN (IN RECF-GYRB INTERGENIC REGION) from Streptomyces coelicolor (190 aa), FASTA scores: opt: 404, E(): 3.9e-18, (40.7% identity in 189 aa overlap)."

```
                     /codon_start=1
                     /transl_table=11
                     /product="CONSERVED HYPOTHETICAL PROTEIN"
                     /protein_id="CAD92866.1"
                     /db_xref="GI:31616766"
                     /translation="MTGSVDRPDQNRGERLMKSPGLDLVRRTLDEARAAARARGQDAG
                     RGRVASVASGRVAGRRRSWSGPGPDIRDPQPLGKAARELAKKRGWSVRVAEGMVLGQW
                     SAVVGHQIAEHARPTALNDGVLSVIAESTAWATQLRIMQAQLLAKIAAAVGNDVVRSL
                     KITGPAAPSWRKGPRHIAGRGPRDTYG"
     gene            5123. .7267
                     /gene="gyrB"
                     /locus_tag="Mb0005"
     CDS             5123. .7267
                     /gene="gyrB"
                     /locus_tag="Mb0005"
                     /EC_number="5.99.1.3"
                     /note="Mb0005, gyrB, len: 714 aa. Equivalent to Rv0005,
                     len: 714 aa, from Mycobacterium tuberculosis strain H37Rv,
                     (99.9% identity in 714 aa overlap). gyrB, DNA gyrase
                     subunit B (EC 5.99.1.3) (see citations below), equivalent,
                     except in N-terminus, to other Mycobacterial DNA GYRASES
                     SUBUNIT B e.g. T10005 from Mycobacterium leprae (697 aa);
                     Q9L7L3|GYRB_MYCPA from Mycobacterium avium subsp.
                     paratuberculosis (677 aa) (has its N-terminus shorter);
                     P48355|GYRB_MYCSM from Mycobacterium smegmatis (675 aa);
                     etc. Also highly similar to others e.g. T10969 from
                     Streptomyces coelicolor (686 aa); P50075|GYBS_STRSH from
```

```
Alignment Scores:
Pred. No.:                1.79e-39        Length:          343050
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       1               Gaps:            0

US-09-684-215B-23 (1-128) x BX248334 (1-343050)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152008 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 152067

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152068 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 152127

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152128 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 152187

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152188 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 152247

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152248 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 152307

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152308 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 152367

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db     152368 ACAGGGAACGTGACATTGGCCGAG 152391

RESULT 13
AR169152
LOCUS       AR169152                 447 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 4 from patent US 6290969.
ACCESSION   AR169152
VERSION     AR169152.1  GI:17906927
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: US 6290969-A 4 18-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                5.52e-42        Length:          447
Score:                    646.00          Matches:         127
Percent Similarity:       99.22%          Conservative:    0
Best Local Similarity:    99.22%          Mismatches:      1
Query Match:              99.23%          Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-23 (1-128) x AR169152 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 14
AR182442
LOCUS       AR182442                 447 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 4 from patent US 6338852.
ACCESSION   AR182442
VERSION     AR182442.1  GI:20225649
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6338852-A 4 15-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                5.52e-42        Length:         447
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-23 (1-128) x AR182442 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 15
AR194825
LOCUS       AR194825                 447 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 4 from patent US 6350456.
ACCESSION   AR194825
VERSION     AR194825.1  GI:20244262
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W. and Dillon,D.C.
  TITLE     Compositions and methods for the prevention and treatment of M.
            tuberculosis infection
  JOURNAL   Patent: US 6350456-A 4 26-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                5.52e-42        Length:         447
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-23 (1-128) x AR194825 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 16
AR233097
LOCUS       AR233097                 447 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 4 from patent US 6458366.
ACCESSION   AR233097
VERSION     AR233097.1  GI:27275533
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6458366-A 4 01-OCT-2002;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                5.52e-42        Length:         447
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-23 (1-128) x AR233097 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
                |||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
                |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 17
AR353302
LOCUS       AR353302                447 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 4 from patent US 6592877.
ACCESSION   AR353302
VERSION     AR353302.1  GI:33759108
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: US 6592877-A 4 15-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                 5.52e-42        Length:          447
Score:                     646.00          Matches:         127
Percent Similarity:        99.22%          Conservative:    0
Best Local Similarity:     99.22%          Mismatches:      1
Query Match:               99.23%          Indels:          0
DB:                        6               Gaps:            0

US-09-684-215B-23 (1-128) x AR353302 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
                |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 18
AX429596
LOCUS       AX429596                447 bp    DNA     linear   PAT 21-JUN-2002
DEFINITION  Sequence 4 from Patent EP1203817.
ACCESSION   AX429596
VERSION     AX429596.1  GI:21540845
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Reed,S.G., Skeiky,Y.A., Dillon,D.C., Campos-Neto,A., Houghton,R.L.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: EP 1203817-A 4 08-MAY-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                 5.52e-42        Length:          447
Score:                     646.00          Matches:         127
Percent Similarity:        99.22%          Conservative:    0
Best Local Similarity:     99.22%          Mismatches:      1
Query Match:               99.23%          Indels:          0
DB:                        6               Gaps:            0

US-09-684-215B-23 (1-128) x AX429596 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
                |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 19
AX832581
LOCUS       AX832581                447 bp    DNA     linear   PAT 12-DEC-2003
DEFINITION  Sequence 4 from Patent EP1347055.
ACCESSION   AX832581
VERSION     AX832581.1  GI:39840631
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Reed,S.G., Skeiky,Y.A., Campos-Neto,A., Houghton,R.L.,
            Vedvick,T.S., Twardzik,D.R. and Dillon,D.C.
```

```
  TITLE     Compounds for immunotherapy and diagnosis of tuberculosis
  JOURNAL   Patent: EP 1347055-A 4 24-SEP-2003;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                5.52e-42        Length:         447
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-23 (1-128) x AX832581 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 20
BD006325
LOCUS       BD006325                447 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Compounds and methods for diagnosis of Tuberculosis.
ACCESSION   BD006325
VERSION     BD006325.1  GI:18634696
KEYWORDS    JP 2001500383-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R. and Lodes,M.J.
  TITLE     Compounds and methods for diagnosis of Tuberculosis
  JOURNAL   Patent: JP 2001500383-A 4 16-JAN-2001;
            CORIXA CORP
COMMENT     OS   Unidentified
            PN   JP 2001500383-A/4
            PD   16-JAN-2001
            PF   07-OCT-1997 JP 1998518432
            PR   11-OCT-1996 US    08/729622,13-MAR-1997 US    08/818111 PI
            STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON, PI   ANTONIO CAMPOS
            NETO,
            PI   RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK, PI
            MICHAEL J LODES
            PC   C12N15/31,C07K14/35,C07K16/12,C12Q1/68,C12N15/62,G01N33/53 CC
             Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                5.52e-42        Length:         447
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-23 (1-128) x BD006325 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 21
BD006445
LOCUS       BD006445                447 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Compounds and methods for immunotherapy and diagnosis of
            Tuberculosis.
ACCESSION   BD006445
VERSION     BD006445.1  GI:18634816
KEYWORDS    JP 2001501832-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R. and Lodes,M.J.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
  JOURNAL   Patent: JP 2001501832-A 4 13-FEB-2001;
            CORIXA CORP
COMMENT     OS   Unidentified
            PN   JP 2001501832-A/4
            PD   13-FEB-2001
            PF   07-OCT-1997 JP 1998518456
            PR   11-OCT-1996 US    08/730510,13-MAR-1997 US    08/818112 PI
```

```
            STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON, PI   ANTONIO CAMPOS
            NETO,
            PI   RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK, PI
            MICHAEL J LODES
            PC   C12N15/31,C07K14/35,A61K39/04,A61K48/00,A61K49/00,C12N15/62,
            PC   C07K19/00,
            PC   G01N33/50,G01N33/60,G01N33/569,C12N1/19,C12N1/20,C12N1/21, PC
             C12N5/10//
            PC   (C12N1/21,C12R1:19)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                5.52e-42       Length:          447
Score:                    646.00         Matches:         127
Percent Similarity:       99.22%         Conservative:    0
Best Local Similarity:    99.22%         Mismatches:      1
Query Match:              99.23%         Indels:          0
DB:                       6              Gaps:            0

US-09-684-215B-23 (1-128) x BD006445 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 22
BD069285
LOCUS       BD069285                 447 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Compounds and methods for immunotherapy and diagnosis of
            tuberculosis.
ACCESSION   BD069285
VERSION     BD069285.1  GI:22614888
KEYWORDS    JP 2001517069-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.H. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
  JOURNAL   Patent: JP 2001517069-A 4 02-OCT-2001;
            CORIXA CORP
COMMENT     OS   Unidentified
            PN   JP 2001517069-A/4
            PD   02-OCT-2001
            PF   30-AUG-1996 JP 1997511464
            PR   01-SEP-1995 US    08/523436,22-SEP-1995 US    08/533634 PR
            22-MAR-1996 US    08/620874,05-JUN-1996 US    08/659683 PR
            12-JUL-1996 US    08/680574
            PI   STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON,ANTONIO CAMPOS
            PI   NETO,
            PI   RAYMOND HOUGHTON,THOMAS H VEDVICK,DANIEL R TWARDZIK PC
            C12N15/31,C07K14/35,A61K38/16,C12N15/62,G01N33/569,C12Q1/68, PC
            C12N5/10,
            PC   C12N1/21//A61K39/04,(C12N1/21,C12R1:19)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Compounds and methods for immunotherapy and diagnosis of  CC
                 tuberculosis
            FH   Key             Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                5.52e-42       Length:          447
Score:                    646.00         Matches:         127
Percent Similarity:       99.22%         Conservative:    0
Best Local Similarity:    99.22%         Mismatches:      1
Query Match:              99.23%         Indels:          0
DB:                       6              Gaps:            0

US-09-684-215B-23 (1-128) x BD069285 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 23
BD205817
LOCUS       BD205817                 447 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Compounds and methods for diagnosis of tuberculosis.
ACCESSION   BD205817
```

```
VERSION     BD205817.1  GI:33015587
KEYWORDS    JP 2002530050-A/4.
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: JP 2002530050-A 4 17-SEP-2002;
            CORIXA CORP
COMMENT     OS   Mycobacterium tuberculosis
            PN   JP 2002530050-A/4
            PD   17-SEP-2002
            PF   17-FEB-1999 JP 2000532132
            PR   18-FEB-1998 US    09/024753,05-MAY-1998 US    09/072596 PI
            STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON,ANTONIO CAMPOS  PI
            NETO,
            PI   RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK,MICHAEL J
            PI   LODES,
            PI   RONALD C HENDRICKSON
            PC   C12N15/09,C07H21/02,C07H21/04,C07K7/00,C07K14/35,C07K16/12, PC
              C07K17/00,
            PC   C07K19/00,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/04,C12Q1/
            PC   68,
            PC   G01N33/569,G01N33/68//A61K38/00,A61K39/04,A61K39/395,A61K39/
            PC   395,A61P31/06,
            PC   (C12N1/21,C12R1:19),C12N15/00,C12N5/00,A61K37/02 CC
            Compounds and methods for diagnosis of tuberculosis. FH   Key
                 Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Mycobacterium tuberculosis'.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Mycobacterium tuberculosis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1773"
ORIGIN

Alignment Scores:
Pred. No.:                5.52e-42        Length:         447
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-23 (1-128) x BD205817 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 24
AR169165
LOCUS       AR169165                1872 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 17 from patent US 6290969.
ACCESSION   AR169165
VERSION     AR169165.1  GI:17906947
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: US 6290969-A 17 18-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                2.33e-41        Length:         1872
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-23 (1-128) x AR169165 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAG 1141

RESULT 25
AR182455
LOCUS       AR182455                1872 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 17 from patent US 6338852.
ACCESSION   AR182455
VERSION     AR182455.1  GI:20225662
KEYWORDS    .
```

```
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6338852-A 17 15-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

Alignment Scores:

Pred. No.:	2.33e-41	Length:	1872
Score:	646.00	Matches:	127
Percent Similarity:	99.22%	Conservative:	0
Best Local Similarity:	99.22%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x AR182455 (1-1872)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAG 1141
```

```
RESULT 26
AR194838
LOCUS       AR194838               1872 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 17 from patent US 6350456.
ACCESSION   AR194838
VERSION     AR194838.1  GI:20244275
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W. and Dillon,D.C.
  TITLE     Compositions and methods for the prevention and treatment of M.
            tuberculosis infection
  JOURNAL   Patent: US 6350456-A 17 26-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

Alignment Scores:

Pred. No.:	2.33e-41	Length:	1872
Score:	646.00	Matches:	127
Percent Similarity:	99.22%	Conservative:	0
Best Local Similarity:	99.22%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x AR194838 (1-1872)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAG 1141
```

```
RESULT 27
AR233110
LOCUS       AR233110               1872 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 17 from patent US 6458366.
ACCESSION   AR233110
VERSION     AR233110.1  GI:27275546
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6458366-A 17 01-OCT-2002;
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

Alignment Scores:

Pred. No.:	2.33e-41	Length:	1872
Score:	646.00	Matches:	127
Percent Similarity:	99.22%	Conservative:	0
Best Local Similarity:	99.22%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x AR233110 (1-1872)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817
```

```
Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy         121 ThrGlyAsnValThrLeuAlaGlu 128
               |||||||||||||||||||||||
Db        1118 ACAGGGAACGTGACATTGGCCGAG 1141
```

RESULT 28
AR353315

LOCUS	AR353315 1872 bp DNA linear PAT 17-AUG-2003
DEFINITION	Sequence 17 from patent US 6592877.
ACCESSION	AR353315
VERSION	AR353315.1 GI:33759121
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown. Unclassified.
REFERENCE	1 (bases 1 to 1872)
AUTHORS	Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R., Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
TITLE	Compounds and methods for immunotherapy and diagnosis of tuberculosis
JOURNAL	Patent: US 6592877-A 17 15-JUL-2003;

FEATURES Location/Qualifiers
source 1. .1872
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.:	2.33e-41	Length:	1872
Score:	646.00	Matches:	127
Percent Similarity:	99.22%	Conservative:	0
Best Local Similarity:	99.22%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x AR353315 (1-1872)

```
Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy         121 ThrGlyAsnValThrLeuAlaGlu 128
               |||||||||||||||||||||||
Db        1118 ACAGGGAACGTGACATTGGCCGAG 1141
```

RESULT 29
AX429609

LOCUS	AX429609 1872 bp DNA linear PAT 21-JUN-2002
DEFINITION	Sequence 17 from Patent EP1203817.
ACCESSION	AX429609
VERSION	AX429609.1 GI:21540858
KEYWORDS	.
SOURCE	unidentified
ORGANISM	unidentified unclassified.
REFERENCE	1
AUTHORS	Reed,S.G., Skeiky,Y.A., Dillon,D.C., Campos-Neto,A., Houghton,R.L., Vedvick,T.S. and Twardzik,D.R.
TITLE	Compounds and methods for immunotherapy and diagnosis of tuberculosis
JOURNAL	Patent: EP 1203817-A 17 08-MAY-2002; CORIXA CORPORATION (US)

FEATURES Location/Qualifiers
source 1. .1872
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:

Pred. No.:	2.33e-41	Length:	1872
Score:	646.00	Matches:	127
Percent Similarity:	99.22%	Conservative:	0
Best Local Similarity:	99.22%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x AX429609 (1-1872)

```
Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy         121 ThrGlyAsnValThrLeuAlaGlu 128
               |||||||||||||||||||||||
Db        1118 ACAGGGAACGTGACATTGGCCGAG 1141
```